



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 160338

TO: Michael Borin
Location: REM-2A55&2C70
Art Unit: 1631
Tuesday, July 26, 2005
Case Serial Number: 10/030605

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

RUSH

THIS PAGE BLANK (USPTO)

160338

From: Chan, Christina
Sent: Monday, July 25, 2005 2:01 PM
To: Borin, Michael; STIC-Biotech/ChemLib
Subject: RE: RUSH search request

Please rush. Thanks Chris

-----Original Message-----

From: Borin, Michael
Sent: Monday, July 25, 2005 1:53 PM
To: Chan, Christina
Subject: RUSH search request

Please approve rush search: case transferred from another examiner, needs additional sequence search

Thank you,

Michael Borin
AU 1631

Examiner: M.Borin
AU: 1631
Mailbox: 2C70
Office: Remsen 2A55
Tel.: 20713

RE: 10030605; crystalline.

Please search of polypeptides SEQ ID 22, 19,21 against commercial and interference protein databases.

Thank you

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:05:40 ; Search time 100.497 Seconds
(without alignments)
758.153 Million cell updates/sec

Title: US-10-030-605A-22
Perfect score: 1113
Sequence: 1 MGKITFYDRFGQHCYECs.....PNSSVSDKLAALBHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003s:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1113	100.0	197	4	Aab46851 Bovine ga
2	1066	95.8	197	4	Aab46850 Bovine ga
3	991	89.0	198	4	Aab46849 Bovine ga
4	944	84.8	198	4	Aab46848 Bovine ga
5	816.5	73.4	173	7	Add47108 Rat Prote
6	816.5	73.4	173	7	Ades7551 Rat Prote
7	816.5	73.4	173	7	Add47114 Rat Prote
8	816.5	73.4	173	7	Add47112 Rat Prote
9	816.5	73.4	173	7	Add47118 Rat Prote
10	816.5	73.4	173	7	Ades8335 Rat Prote
11	816.5	73.4	173	7	Ades7555 Rat Prote
12	788.5	70.8	173	7	Add47104 Rat Prote
13	758.5	68.1	174	4	Abg21006 Novel hum
14	758.5	68.1	174	7	Add47106 Human pro
15	753.5	67.7	362	7	Adc31213 Human nov
16	749.5	67.3	173	7	Add47110 Human pro
17	749.5	67.3	173	7	Ades83337 Human pro
18	749.5	67.3	173	7	Add47116 Human pro
19	749.5	67.3	173	7	Ades7553 Human pro
20	732	65.8	308	4	Abg21004 Novel hum
21	707.5	63.6	199	4	Abg21007 Novel hum
22	635	57.1	145	4	Abg21005 Novel hum
23	345	31.0	169	8	AQg97976 Human can
24	331	29.7	781	5	Abb72344 Murine pr
25	328	29.5	507	7	Adc31503 Human nov

26	328	29.5	616	7	ADM04410 Human pro
27	328	29.5	888	8	ADO44168 Structura
28	327	29.4	211	8	ADP22673 Golden ha
29	323.5	29.1	420	5	ABP69298 Human pol
30	319	28.7	204	7	ADP63507 Rat Prote
31	319	28.7	205	8	ADP22671 Golden ha
32	314	28.2	205	2	AAW17523 Human bet
33	309.5	27.8	215	2	AAW17522 Human bet
34	302	27.1	252	5	AAU11447 Human cry
35	285.5	25.7	542	6	ABR41645 Human DIT
36	285.5	25.7	962	7	ADP29065 Human ATM
37	285.5	25.7	1723	6	ABG74680 Human CGD
38	284	25.5	250	8	ADP22669 Golden ha
39	283.5	25.5	1080	4	AAW00803 Human bon
40	283.5	25.5	1637	4	AAW00916 Human bon
41	283.5	25.5	2263	4	AAW79000 Human pro
42	273	24.5	97	3	AB58947 Breatst an
43	272	24.4	197	4	AB471135 CDIPP-16,
44	272	24.4	197	4	ABG08246 Novel hum
45	272	24.4	197	8	ADQ19127 Human sof

ALIGNMENTS

RESULT 1
AAB46851
ID AAB46851 standard; protein; 197 AA.
XX
AC AAB46851;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline protein WT-HIS fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW Biosensor; pollution detection; pollution control; gene therapy;
KW Intracellular immunization.
XX
OS Bos taurus.
XX
FN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 95DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FTED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 19-20; 28pp; German.
XX
CC This invention describes a novel protein (I) with beta-'leaflet',
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (i) DNA (ii) that encodes (i); (2) RNA (iii)
CC derived from (ii); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (ii) or (iii), or their fragments that encode a functional region
CC of (i); and (4) method for producing (i). (i) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene

CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;

Query Match 100.0%; Score 1113; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 8.6e-111;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGKITFYEDRGQGHCHYCSCDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
DB 1 MGKITFYEDRGQGHCHYCSCDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
QY 61 GDYPDYQQWGMGNDIRSCLIPQHTGTFRMRYIERDDFRGOMSEITDDCPSLQDRFHLT 120
DB 61 GDYPDYQQWGMGNDIRSCLIPQHTGTFRMRYIERDDFRGOMSEITDDCPSLQDRFHLT 120
QY 121 EVHSLNVLEGSWLVYEMPSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
DB 121 EVHSLNVLEGSWLVYEMPSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
QY 181 SSVDKLAAALEHHHHHH 197
DB 181 SSVDKLAAALEHHHHHH 197

RESULT 2
AAB46850
ID AAB46850 standard; protein; 197 AA.
XX
AC AAB46850;
DT 26-APR-2001 (first entry)
DE Bovine gamma-crystalline mutant protein Mul2A-His fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
XX WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 18-19; 28pp; German.

CC This invention describes a novel protein (I) with beta-'leaflet'
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The

CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene
CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;

Query Match 95.8%; Score 1066; DB 4; Length 197;
Best Local Similarity 96.4%; Pred. No. 9.5e-106;
Matches 190; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 MGKITFYEDRGQGHCHYCSCDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
DB 1 MGKITFYEDRGQGHCHYCSCDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
QY 61 GDYPDYQQWGMGNDIRSCLIPQHTGTFRMRYIERDDFRGOMSEITDDCPSLQDRFHLT 120
DB 61 GDYPDYQQWGMGNDIRSCLIPQHTGTFRMRYIERDDFRGOMSEITDDCPSLQDRFHLT 120
QY 121 EVHSLNVLEGSWLVYEMPSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
DB 121 EVHSLNVLEGSWLVYEMPSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
QY 181 SSVDKLAAALEHHHHHH 197
DB 181 SSVDKLAAALEHHHHHH 197

RESULT 3
AAB46849
ID AAB46849 standard; protein; 198 AA.
XX
AC AAB46849;
DT 26-APR-2001 (first entry)
DE Bovine gamma-crystalline protein fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
XX WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 17-18; 28pp; German.

XX This invention describes a novel protein (I) with beta-'leaflet',
 CC structure having surface-exposed amino acids, present in at least two
 CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
 CC protein is altered by targeted mutagenesis so that it has new, or
 CC improved, specific binding, catalytic or fluorescent properties. The
 CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
 CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
 CC contain (II) or (III), or their fragments that encode a functional region
 CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
 CC and therapy, in cosmetics, bioseparation and biosensors, and for
 CC pollution detection and control, e.g. for specific targeting of gene
 CC therapy vectors and for intracellular immunization. (I) can be provided
 CC with new or improved specific antibody-like binding, catalytic or
 CC fluorescent properties, without the cost and difficulties associated with
 CC producing complete or recombinant antibodies. (I) are relatively small
 CC (20 kDa) and can be expressed with other components as multifunctional
 CC fusions. They have good stability against low pH, denaturing agents and
 CC high temperatures, conditions under which antibodies are unstable
 XX Sequence 198 AA;

Query Match 89.0%; Score 991; DB 4; Length 198;
 Best Local Similarity 99.4%; Pred. No. 1.1e-97;
 Matches 175; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKITYEDRGFGHCHYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
 DB 6 MGKITYEDRGFGHCHYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 65
 QY 61 GDYPDYQWGMFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSBITDDCPSLQDRFHLT 120
 DB 66 GDYPDYQWGMFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSBITDDCPSLQDRFHLT 125
 QY 121 EVHSLNVLEGSWLVEMPSYRGRQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 176
 DB 126 EVHSLNVLEGSWLVEMPSYRGRQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 181

RESULT 4
 ID AAB46848 standard; protein; 198 AA.
 AC AAB46848;
 XX 26-APR-2001 (first entry)
 XX Bovine gamma-crystalline mutant Mul2A protein fragment.
 XX Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
 KW biosensor; pollution detection; pollution control; gene therapy;
 KW intracellular immunization.
 XX Bos taurus.
 OS DE19932688-A1.
 XX 18-JAN-2001.
 XX 13-JUL-1999; 99DE-01032688.
 XX 13-JUL-1999; 99DE-01032688.
 XX (FIEDLER) FIEDLER U.
 XX (RUDOLPH) RUDOLPH R.
 XX Rudolph R, Fiedler U, Boehm G, Reimann C;
 XX WPI; 2001-148304/16.
 XX Mutated proteins having beta-leaflet structure and related nucleic acid,
 PT have new or improved properties, e.g. antibody-like specific binding or
 PT catalytic activity.

XX Claim 11; Page 16-17; 28pp; German.
 PS This invention describes a novel protein (I) with beta-'leaflet',
 CC structure having surface-exposed amino acids, present in at least two
 CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
 CC protein is altered by targeted mutagenesis so that it has new, or
 CC improved, specific binding, catalytic or fluorescent properties. The
 CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
 CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
 CC contain (II) or (III), or their fragments that encode a functional region
 CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
 CC and therapy, in cosmetics, bioseparation and biosensors, and for
 CC pollution detection and control, e.g. for specific targeting of gene
 CC therapy vectors and for intracellular immunization. (I) can be provided
 CC with new or improved specific antibody-like binding, catalytic or
 CC fluorescent properties, without the cost and difficulties associated with
 CC producing complete or recombinant antibodies. (I) are relatively small
 CC (20 kDa) and can be expressed with other components as multifunctional
 CC fusions. They have good stability against low pH, denaturing agents and
 CC high temperatures, conditions under which antibodies are unstable
 XX Sequence 198 AA;

Query Match 84.8%; Score 944; DB 4; Length 198;
 Best Local Similarity 95.5%; Pred. No. 1.2e-92;
 Matches 168; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGKITYEDRGFGHCHYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
 DB 6 MGKITYEDRGFGHCHYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 65
 QY 61 GDYPDYQWGMFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSBITDDCPSLQDRFHLT 120
 DB 66 GDYPDYQWGMFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSBITDDCPSLQDRFHLT 125
 QY 121 EVHSLNVLEGSWLVEMPSYRGRQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 176
 DB 126 EVHSLNVLEGSWLVEMPSYRGRQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 181

RESULT 5
 ID ADD47108 standard; protein; 173 AA.
 AC ADD47108;
 XX 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)
 XX Rat Protein P10065, SEQ ID NO 12798.
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX Rattus norvegicus.
 OS Unidentified.
 XX WO20003016475-A2.
 XX 27-FEB-2003.
 XX 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX

DR WPI; 2003-268312/26.
XX GENBANK; P10065.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX
PS Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 173 AA;
SQ
Query Match 73.4%; Score 816.5; DB 7; Length 173;
Best Local Similarity 81.6%; Pred. No. 4.7e-79;
Matches 142; Conservative 18; Mismatches 13; Indels 1; Gaps 1;
QY 2 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERNYQGHQYFLRRG 61
DB 1 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERNYQGHQYFLRRG 60
QY 62 DYPDYQQWGMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTDCSCIHDRFRLNE 121
DB 61 DYPDYQQWGMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTDCSCIHDRFRLNE 119
QY 122 VHSLNVLGSGWLYEMPSYRGROYLLRCEYRRLDGMGMNAKVGSLRRVMDFY 175
DB 120 IYSMHVLEGSWLYEMPNYRGROYLLRCEYRRLDGMGMNAKVGSLRRVMDLY 173
RESULT 6
ADES7551
ID ADES7551 standard; protein; 173 AA.
XX
XX ADES7551;
AC
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P10065, SEQ ID NO 3413.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.
FN
XX
XX 27-FEB-2003.
PD

XX
PF 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
XX GENBANK; P10065.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 173 AA;
SQ
Query Match 73.4%; Score 816.5; DB 7; Length 173;
Best Local Similarity 81.6%; Pred. No. 4.7e-79;
Matches 142; Conservative 18; Mismatches 13; Indels 1; Gaps 1;
QY 2 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERNYQGHQYFLRRG 61
DB 1 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERNYQGHQYFLRRG 60
QY 62 DYPDYQQWGMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTDCSCIHDRFRLNE 121
DB 61 DYPDYQQWGMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTDCSCIHDRFRLNE 119
QY 122 VHSLNVLGSGWLYEMPSYRGROYLLRCEYRRLDGMGMNAKVGSLRRVMDFY 175
DB 120 IYSMHVLEGSWLYEMPNYRGROYLLRCEYRRLDGMGMNAKVGSLRRVMDLY 173
RESULT 7
ADD47114
ID ADD47114 standard; protein; 173 AA.
XX
XX AC
XX ADD47114;
XX

```
DT 02-DEC-2004 (revised)
XX 29-JAN-2004 (first entry)
DE Rat Protein P10065, SEQ ID NO 12804.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
OS Unidentified.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; P10065.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (described in Table 3
XX of the specification) which is differentially expressed during pain.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 173 AA;
XX
XX Query Match 73.4%; Score 816.5; DB 7; Length 173;
XX Best Local Similarity 81.6%; Pred. No. 4.7e-79;
XX Matches 142; Conservative 18; Mismatches 13; Indels 1; Gaps 1;
XX
XX 2 GKITYEDRGFGCHYCECSDCPNLQYFSCRNSRVDSCWMLYERPNYQGHQYFLRRG 61
XX
XX 1 GKITYEDRGFGORCEYCSDDCPNLQYFSCRNSRVDSCWMLYERPNYQYQYFLRRG 60
XX
XX 62 DYPDYQQWGMFNDNSRSCRLIPQHTGTFRMRYERDDFRGQMSBITDDCPSLQDRFHLTE 121
XX
```

```
Db 61 DYPDYQQWGMFSDSIRSRSIP-YTSSHRIRLYERDDYRGLVSELTDCSCIHDRFLNE 119
QY 122 VHSLNVLGSGSWLYEMPYRGQYLLRPGYRRYLLDGMAMNAKVGSLRRVMDFY 175
Db 120 IYSMHVLESGSWLYEMPYRGQYLLRPGYRRYLLDGMAMNAKVGSLRRVMDLY 173
RESULT 8
ADD47112
ID ADD47112 standard; protein; 173 AA.
XX
XX AC ADD47112;
XX
XX DT 02-DEC-2004 (revised)
XX DT 29-JAN-2004 (first entry)
XX
XX DE Rat Protein AAA40981, SEQ ID NO 12802.
XX
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX OS Rattus norvegicus.
XX OS Unidentified.
XX
XX PN WO2003016475-A2.
XX
XX PD 27-FEB-2003.
XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; AAA40981.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (described in Table 3
XX of the specification) which is differentially expressed during pain.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 173 AA;
XX
```

CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 173 AA;

Query Match 73.4%; Score 816.5; DB 7; Length 173;
Best Local Similarity 81.6%; Pred. No. 4.7e-79;
Matches 142; Conservative 18; Mismatches 13; Indels 1; Gaps 1;

QY 2 GKITYEDRGFGCHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 61
|||||
Db 1 GKITYEDRGFGCHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 60
|||||

QY 62 DYDPYQQWGMFNDIRSCLIPQHTGTFRMRYERDDRGOMSEITDDCPSLQDRFHLTE 121
|||||
Db 61 DYDPYQQWGMFNDIRSCLIPQHTGTFRMRYERDDRGOMSEITDDCPSLQDRFHLTE 119
|||||

QY 122 VHSNLVLEGSWLYEMPYSGRQYLLRPGCYRRYLDWGAMNAKVGLSLRRVMDFY 175
:::|||||
Db 120 IYSMHVLEGSWLYEMPYSGRQYLLRPGCYRRYLDWGAMNAKVGLSLRRVMDLY 173
|||||

RESULT 9
ADD47118
ID ADD47118 standard; protein; 173 AA.
XX
AC ADD47118;
XX
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX

DE Rat Protein AAA40981, SEQ ID NO 12808.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
OS Unidentified.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; AAA40981.
XX

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Example 1; Page; 1017pp; English.
XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 173 AA;

Query Match 73.4%; Score 816.5; DB 7; Length 173;
Best Local Similarity 81.6%; Pred. No. 4.7e-79;
Matches 142; Conservative 18; Mismatches 13; Indels 1; Gaps 1;

QY 2 GKITYEDRGFGCHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 61
|||||
Db 1 GKITYEDRGFGCHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 60
|||||

QY 62 DYDPYQQWGMFNDIRSCLIPQHTGTFRMRYERDDRGOMSEITDDCPSLQDRFHLTE 121
|||||
Db 61 DYDPYQQWGMFNDIRSCLIPQHTGTFRMRYERDDRGOMSEITDDCPSLQDRFHLTE 119
|||||

QY 122 VHSNLVLEGSWLYEMPYSGRQYLLRPGCYRRYLDWGAMNAKVGLSLRRVMDFY 175
:::|||||
Db 120 IYSMHVLEGSWLYEMPYSGRQYLLRPGCYRRYLDWGAMNAKVGLSLRRVMDLY 173
|||||

RESULT 10
ADE83335
ID ADE83335 standard; protein; 173 AA.
XX
AC ADE83335;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P10065, SEQ ID NO 10928.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P10065.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 173 AA;

Query Match 73.4%; Score 816.5; DB 7; Length 173;
 Best Local Similarity 81.6%; Pred. No. 4.7e-79;
 Matches 142; Conservative 18; Mismatches 13; Indels 1; Gaps 1;
 QY 2 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 61
 DB 1 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 60
 QY 62 DYPDYQQWGMFSDSIRSCRSIP-YTSSHRIYERDDYRGLVSELTDCSCIHDRFRNLN 121
 DB 61 DYPDYQQWGMFSDSIRSCRSIP-YTSSHRIYERDDYRGLVSELTDCSCIHDRFRNLN 119
 QY 122 VHSNLVLEGSWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDFY 175
 DB 120 IYSMHVLEGSWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDLY 173

RESULT 11
 ADE57555
 ID ADE57555 standard; protein; 173 AA.
 AC ADE57555;
 XX 29-JAN-2004 (first entry)
 DT Rat Protein AAA40981, SEQ ID NO 3417.
 DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX Rattus norvegicus.
 OS WO2003016475-A2.
 PN 27-FEB-2003.
 PD 14-AUG-2002; 2002WO-05025765.
 PF 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX

PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-269312/26.
 DR GENBANK; AAA40981.
 XX

New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 173 AA;

Query Match 73.4%; Score 816.5; DB 7; Length 173;
 Best Local Similarity 81.6%; Pred. No. 4.7e-79;
 Matches 142; Conservative 18; Mismatches 13; Indels 1; Gaps 1;
 QY 2 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 61
 DB 1 GKITYEDRGFGQHCYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRRG 60
 QY 62 DYPDYQQWGMFSDSIRSCRSIP-YTSSHRIYERDDYRGLVSELTDCSCIHDRFRNLN 121
 DB 61 DYPDYQQWGMFSDSIRSCRSIP-YTSSHRIYERDDYRGLVSELTDCSCIHDRFRNLN 119
 QY 122 VHSNLVLEGSWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDFY 175
 DB 120 IYSMHVLEGSWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDLY 173

RESULT 12
 ADD47104
 ID ADD47104 standard; protein; 173 AA.
 XX ADD47104;
 AC ADD47104;
 XX 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)
 DE Rat Protein AAA40988, SEQ ID NO 12794.
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

Db 61 GDYADHQWGLSDSVSRCLIP-HSGSHRILRYERDVRGQMIETEDCSCLQDRFRFN 119
QY 121 EVHSLNVLEGSWLYEMPSYGRQYLLRPGYRRYLDGAMNAKVGSLRVNDF 174
Db 120 EIHSNLVLEGSWLYELSNYGRQYLLMPGDYRRYQDWGATNARVGSLLRRVIDF 173

RESULT 14

ID ADD47106 standard; protein; 174 AA.

AC ADD47106;

XX 02-DEC-2004 (revised)

DT 29-JAN-2004 (first entry)

XX Human Protein NP_008822, SEQ ID NO 12796.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

OS Unidentified.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI: 2003-268312/26.

DR GENPANK; NP_008822.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 174 AA;

Query Match 68.1%; Score 758.5; DB 7; Length 174;

Best Local Similarity 77.0%; Pred. No. 8e-73; Mismatches 17; Indels 1; Gaps 1;
Matches 134; Conservative 17;

QY 1 MGKITFYEDRGFGQHCYECSSDCPNLQPYFVSRCSNIRVDSGCWMLYERPNYQGHVFLRR 60

Db 1 MGKITLYEDRGFGQHRHYECSSDHPNLPYLSCNARSARVDSGCWMLYEQNYSGQLYFLRR 60

QY 61 GDYDYQWQMGFNDSIRSCRLIPQHTGTFRRIYERDDFRGQMSITDDCPSLQDRFHLT 120

Db 61 GDYADHQWGLSDSVSRCLIP-HSGSHRILRYERDVRGQMIETEDCSCLQDRFRFN 119

QY 121 EVHSLNVLEGSWLYEMPSYGRQYLLRPGYRRYLDGAMNAKVGSLRVNDF 174

Db 120 EIHSNLVLEGSWLYELSNYGRQYLLMPGDYRRYQDWGATNARVGSLLRRVIDF 173

RESULT 15

ADC31213

ID ADC31213 standard; protein; 362 AA.

AC ADC31213;

DT 18-DEC-2003 (first entry)

XX Human novel polypeptide sequence, SEQ ID NO:1295.

XX Human; diagnostic; drug screening; forensic; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; Burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 18.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QH, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;

PI Haley-Vicente D, Drmanac RT;

DR WPI: 2003-371981/35.

DR N-PSDB; ADC30242.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.

XX Claim 20; SEQ ID NO 1295; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:18:55 ; Search time 25.9561 Seconds
(without alignments)
566.567 Million cell updates/sec

Title: US-10-030-605A-22
Perfect score: 1113
Sequence: 1 MGKITYEDRFGQHCYECS.....PNSSVDKLAALHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	860	77.3	187	4	US-09-949-016-9254
2	758.5	68.1	179	4	US-09-949-016-11310
3	555.5	49.9	192	4	US-09-949-016-11340
4	322.5	29.0	204	4	US-09-949-016-9696
5	314	28.2	205	2	US-08-729-152-8
6	313	28.1	235	4	US-09-949-016-8099
7	310	27.9	209	4	US-09-949-016-10830
8	309.5	27.8	215	2	US-08-729-152-1
9	302	27.1	262	4	US-09-949-016-10815
10	294.5	26.5	207	4	US-09-949-016-10814
11	100	9.0	42	1	US-08-664-449-39
12	86	7.7	378	4	US-08-979-847B-122
13	86	7.7	398	4	US-08-979-847B-121
14	84.5	7.6	1010	4	US-09-654-449-2
15	84.5	7.6	1010	4	US-09-759-152A-2
16	82	7.4	320	2	US-08-757-653-163
17	82	7.4	320	2	US-08-823-516-61
18	82	7.4	320	3	US-08-759-038-102
19	82	7.4	320	3	US-08-758-314-102
20	82	7.4	320	4	US-09-684-938-102
21	82	7.4	320	4	US-09-308-825A-102
22	82	7.4	320	4	US-09-940-244-61
23	80.5	7.2	243	4	US-09-543-681A-7356
24	80	7.2	3854	4	US-09-949-016-7876
25	79.5	7.1	139	4	US-09-621-976-5088
26	79.5	7.1	300	3	US-09-561-756-36
27	79.5	7.1	300	3	US-09-227-721-36

28	79.5	7.1	300	4	US-09-954-697-36	Sequence 36, Appl
29	79.5	7.1	315	2	US-08-484-956-91	Sequence 91, Appl
30	79.5	7.1	315	2	US-08-757-653-91	Sequence 91, Appl
31	79.5	7.1	315	3	US-08-520-946-91	Sequence 91, Appl
32	79.5	7.1	315	4	US-09-655-378A-91	Sequence 91, Appl
33	77	6.9	225	4	US-09-456-090A-94	Sequence 94, Appl
34	77	6.9	225	4	US-09-453-234-94	Sequence 94, Appl
35	76	6.8	229	4	US-09-546-043-8	Sequence 8, Appl
36	76	6.8	338	4	US-08-816-977-21	Sequence 21, Appl
37	75.5	6.8	323	3	US-09-334-477-21	Sequence 21, Appl
38	75.5	6.8	323	4	US-09-270-767-43477	Sequence 43477, A
39	75.5	6.8	655	4	US-09-489-039A-11003	Sequence 11003, A
40	75.5	6.8	825	4	US-09-456-090A-58	Sequence 58, Appl
41	75	6.7	225	4	US-09-453-234-58	Sequence 9, Appl
42	75	6.7	225	4	US-09-387-375-9	Sequence 9, Appl
43	74.5	6.7	316	4	US-10-041-400A-9	Sequence 9, Appl
44	74.5	6.7	316	4	US-10-042-091A-9	Sequence 9, Appl
45	74.5	6.7	316	4	US-10-042-091A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-949-016-9254
; Sequence 9254, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 20702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9254
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9254

Query Match	77.3%	Score	860	DB	4	Length	187
Best Local Similarity	84.0%	Pred. No.	1.5e-85				
Matches	147	Conservative	17	Mismatches	11	Indels	0
Gaps	0						
QY	1	MGKITYEDRFGQHCYECS	DCNLPQYFSCN	SIRVDSGCMWLYERNYQGHYELRR	60		
DB	13	MGKITYEDRFGQHCYECS	DCNLPQYFSCN	SIRVDSGCMWLYERNYQGHYELRR	72		
QY	61	GDYDPYQOMGPNDSIRSCRLIPQHTGTFMRIRYRDDPRGOMSEITDDCPSLQDRFHLT	120				
DB	73	GEYDPYQOMGPNDSIRSCRLIPQHTGTFMRIRYRDDPRGOMSEITDDCPSLQDRFHLT	132				
QY	121	EVHSLNVLGSGWLYEMPSYGRQYLLRPGEYRRYLDGMGNNAKVGSLRRVMDFY	175				
DB	133	BIHSLNVLGSGWLYEMPSYGRQYLLRPGEYRRYLDGMGNNAKVGSLRRVMDFY	187				

RESULT 2
US-09-949-016-11310
; Sequence 11310, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11310
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11310

Query Match      68.1%; Score 758.5; DB 4; Length 179;
Best Local Similarity 77.0%; Pred. No. 1.5e-74;
Matches 134; Conservative 17; Mismatches 22; Indels 1; Gaps 1;

QY 1 MCKITFYEDRGFQGHYECSSDCPNLQPYFSCNIRVDSCGMWLYERPNTYQGHQYFLRR 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 MCKITFYEDRGFQGHYECSSDCPNLQPYFSCNIRVDSCGMWLYEQPNYSGLYFLRR 65
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GYDPYQQWGMFNDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRFHLT 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 GYADHQQWGLSDSVRSCLLIP-HSGSHRIKLYERDYGQMIEFTDCSLQDRFRFN 124
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 EVHSLNVLEGSWLVLEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGLRRVMD 174
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 EIHSLNVLEGSWLVLESNYRGQYLLMPGDYRRYQDGNATNARVGLRRVIDF 178
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-09-949-016-11340
; Sequence 11340, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11340
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11340

Query Match      49.9%; Score 555.5; DB 4; Length 192;
Best Local Similarity 52.9%; Pred. No. 2.1e-52;
Matches 91; Conservative 36; Mismatches 44; Indels 1; Gaps 1;

QY 3 KITFYEDRGFQGHYECSSDCPNLQPYFSCNIRVDSCGMWLYERPNTYQGHQYFLRRGD 62
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 KITFYEDKNFGRRYDCDCADFTYLSRCNSIKVEGTVAVYERPNTYQGHQYFLRRGD 80
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 63 YDPYQQWGMFNDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRFHLT 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 YPEYQRWGLNDRLSRCRAVHLPSSGGYKIQIFKGFSGQNYETTEDCPSLQDRFHLT 140
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 122 VHSLSNVLEGSWLVLEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGLRRVMD 173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 291993/1995
FILING DATE: 13-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 205 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLSCULE TYPE: protein

DESCRIPTION: protein

US-08-729-152-8

Query Match 28.2%; Score 314; DB 2; Length 205;

Best Local Similarity 37.1%; Pred. No. 4.2e-26;

Matches 65; Conservative 33; Mismatches 73; Indels 4; Gaps 3;

QY 3 KITFYEDRGFGCHCYECSSDCPNL-QPYFSCNSIRVDSGCMWLYERPNYQGHQYFLRRG 61

DB 18 KIIIFEQENFQGHSHELNGPCPNLKETGVKAGSVLVQAGPWVGYEQANCKGEQVFEKG 77

QY 62 DYPDYQWQMGF--NDSIRSCRLLPQHTGTFRMRYERDDFRQMSSE-ITDDCPSLQDRPH 118

DB 78 EYPRWDSWTSSRRDTSLSLRPIKVDSEHKKIILYENPNFTGKMEIIDDVPSFHAHY 137

QY 119 LTVHSLNVLGSGWLYEMPSYRGQYLLRPGYRRYLDGAMNAKVGLRRVMD 173

DB 138 QEKVSSVVRSGTWGQYQYGRGLQYLLEKGDYKDSDFGAPHPQVQSVRRIRD 192

RESULT 6

US-09-949-016-8099

Sequence 8099, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8099

LENGTH: 235

TYPE: PRT

ORGANISM: Human

US-09-949-016-8099

Query Match 28.1%; Score 313; DB 4; Length 235;

Best Local Similarity 37.2%; Pred. No. 6.5e-26;

Matches 67; Conservative 33; Mismatches 66; Indels 14; Gaps 5;

QY 3 KITFYEDRGFGCHCYECSSDCPNL-QPYFSCNSIRVDSGCMWLYERPNYQGHQYFLRRG 61

DB 49 KVIYELNFGKRCLESAECPSLDLSLEKVGSLQVESGFWLAFESAFRGEQVFEKG 108

QY 62 DYPDYQWQMGF--NDSIRSCRLLPQHTGTFRMRYERDDFRQMSSEIT-DDCPSL----- 113

DB 109 DYPRWDAMSGRDSDLSLSLRPLNIDSPDHKLHLFENPAFSGRKMIEIVDDDDVPSLWAHGF 168

QY 114 QDRFHLTEVHSLNVLGSGWLYEMPSYRGQYLLRPGYRRYLDGAMNAKVGLRRVMD 173

DB 169 QDR-----VASVRAINGTWGVEYFPGYRGQYVPERGEGYRHNWEDASQPQLQSVRRIRD 223

RESULT 7

US-09-949-016-10830

Sequence 10830, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10830

LENGTH: 209

TYPE: PRT

ORGANISM: Human

US-09-949-016-10830

Query Match 27.9%; Score 310; DB 4; Length 209;

Best Local Similarity 37.1%; Pred. No. 1.2e-25;

Matches 65; Conservative 33; Mismatches 73; Indels 4; Gaps 3;

QY 3 KITFYEDRGFGCHCYECSSDCPNL-QPYFSCNSIRVDSGCMWLYERPNYQGHQYFLRRG 61

DB 22 KIIIFEQENFQGHSHELNGPCPNLKETGVKAGSVLVQAGPWVGYEQANCKGEQVFEKG 81

QY 62 DYPDYQWQMGF--NDSIRSCRLLPQHTGTFRMRYERDDFRQMSSE-ITDDCPSLQDRPH 118

DB 82 EYPRWDSWTSSRRDTSLSLRPIKVDSEHKKIILYENPNFTGKMEIIDDVPSFHAHY 141

QY 119 LTVHSLNVLGSGWLYEMPSYRGQYLLRPGYRRYLDGAMNAKVGLRRVMD 173

DB 142 QEKVSSVVRSGTWGQYQYGRGLQYLLEKGDYKDSDFGAPHPQVQSVRRIRD 196

RESULT 8

US-08-729-152-1

Sequence 1, Application US/08729152

Patent No. 5871739

GENERAL INFORMATION:

APPLICANT: Inoue, Eri

TITLE OF INVENTION: Pharmaceutical Composition

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack

STREET: Southern Building, Suite 700, 805 Fifteenth

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729,152

```
;
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 291993/1995
; FILING DATE: 13-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: protein
;
US-08-729-152-1
Query Match 27.8%; Score 309.5; DB 2; Length 215;
Best Local Similarity 36.3%; Pred. No. 1.4e-25;
Matches 66; Conservative 38; Mismatches 65; Indels 13; Gaps 6;

QY 3 KITFVEDRGFQGHCEVSSDCPNLQPY-FSRCSNIRVDSCGMWLYERNYQGHQYFLRRG 61
Db 70 RLTVFELENFQGRRAEFGCECSNLRADRGFDRVRSIIVSAGPWAFEQSNFRGEMFILSKG 129
QY 62 DYPDYQQWGMGF--NDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSSEIT-DPCPSLQDRFH 118
Db 130 EYPRNWTWSSSYRSRDLMSFRPIKMDAQEHKISLFGANFKGNTIEIQGDDAPSLWVYGF 189
QY 119 LTEVHSLNVLEGSWLYEMPSYRGROYLLRGEYRRLYLDGAMNAKVGSLLRRVMD 173
Db 190 SDRVGSVKVSSGTWVGYPGYRGYQYLLPEGDFRHNWEGAFQFQMOSLRRLRD 244

RESULT 10
US-09-949-016-10814
; Sequence 10814, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10814
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10814
Query Match 26.5%; Score 294.5; DB 4; Length 207;
Best Local Similarity 35.9%; Pred. No. 5.6e-24;
Matches 66; Conservative 40; Mismatches 61; Indels 17; Gaps 7;

QY 3 KITFVEDRGFQGHCEVSSDCPN-LQPYFSRCSNIRVDSCGMWLYERNYQGHQYFLRRG 61
Db 24 KMWVDEDFGFORRHEFTAECPVLELGFETVRSKLVLSGAWVGFEGHAGFQGGYILERG 83
QY 62 DYPDYQQWGMFNDIRSRL-----IPQHTGTFRMRIYERDDFRGQMSSEITDDCPSLQ 114
Db 84 EYPSWDAM-GGNTATYPAERLTSFRPAACANHRDS-RLTTFEQENFLGKKGLSDDYPSLQ 141
QY 115 DR-PHLTEVHSLNVLEGSWLYEMPSYRGROYLL----RPGEYRRLYLDWG--AMNAKVGS 167
Db 142 AMGVEGNEVSGFVHSHGAWVCSQFPGYRGYQYVLECDHSHGDKYHFRWEGSHAPTFQVQS 201
QY 168 LRRV 171
Db 202 IRRV 205

RESULT 11
US-08-664-449-39
; Sequence 39, Application US/08664449
; Patent No. 5766905
; GENERAL INFORMATION:
; APPLICANT: Studier, F. W.
; APPLICANT: Rosenberg, Alan H.
; TITLE OF INVENTION: Cytoplasmic Bacteriophage Display System
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET: Building 902C
; CITY: Upton
; STATE: NY
```

```
;
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 291993/1995
; FILING DATE: 13-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: protein
;
US-08-729-152-1
Query Match 27.8%; Score 309.5; DB 2; Length 215;
Best Local Similarity 36.3%; Pred. No. 1.4e-25;
Matches 66; Conservative 38; Mismatches 65; Indels 13; Gaps 6;

QY 3 KITFVEDRGFQGHCEVSSDCPNL-QPYFSRCSNIRVDSCGMWLYERNYQGHQYFLRRG 61
Db 32 KITIYDQENFQGRMEFTSSCPNVSERSFDNRSKLVESGAWIGYEHTSFCCQQFILERG 91
QY 62 DYPDYQQWGMGF--NDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSSEITDDCPSLQ 116
Db 92 EYPRNWTWSSSYRSRDLMSFRPIKMDAQEHKISLFGANFKGNTIEIQGDDAPSLWVYGF 189
QY 117 -PHLTEVHSLNVLEGSWLYEMPSYRGROYLLR---PGEYRRLYLDWG--AMNAKVGSILR 169
Db 152 GWFNNEVSGMKIQSGAWCYHYLGRGYQYLLKCDHHEGDKYHFRWEGSHAPTFQVQS 211
QY 170 RV 171
Db 212 RI 213

RESULT 9
US-09-949-016-10815
; Sequence 10815, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10815
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10815
Query Match 27.1%; Score 302; DB 4; Length 262;
Best Local Similarity 36.6%; Pred. No. 1.2e-24;
Matches 64; Conservative 32; Mismatches 75; Indels 4; Gaps 3;
```

COUNTRY: US
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/664,449
FILING DATE: 17-June-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: AUI-9618
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 344-7338
TELEFAX: (516) 344-3729
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-664-449-39

```

Query Match      9.0%; Score 100; DB 1; Length 42;
Best Local Similarity 53.8%; Pred. No. 0.0008;
Matches 21; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

QY      177 DPNSSSVDKLAAAL-----EHHHHH 197
          |||||
Db       4  DPNSSSVDKLAAAL EIKRASOPELAPEDVEHHHHH 42
          |||||

```

RESULT 12
US-08-979-847B-122
; Sequence 122, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A

```

,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: 703-836-6400
,
, TELEFAX: 703-836-2787
,
, INFORMATION FOR SEQ ID NO: 132:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 378 amino acids
,
, TYPE: amino acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: peptide
,
, SEQUENCE DESCRIPTION: SEQ ID NO: 132:
US-08-979-847B-122

```

```

Query Match          7.7%; Score 86; DB 4; Length 378;
Best Local Similarity 78.9%; Pred. No. 0.63;
Matches 15; Conservative 1; Mismatches 3; Indels

QY      179  NSSVDKLAALALEHHHHH 197
      |  -  |||||
Db       360  NFKSLPKLAALALEHHHHH 378

```

RESULT 13
US-08-979-847B-121
; Sequence 121, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

Search completed: July 25, 2005, 16:39:17
Job time : 26.9561 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:36:56 ; Search time 89.5152 Seconds

(without alignments)

856.072 Million cell updates/sec

Title: US-10-030-605A-22

Perfect score: 1113

Sequence: 1 MGKTFYEDRGFGQHCYECSSDPCNLQYFRCNSIRVDSGCWMLYERPNYQGHQYFLRR 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1741741 seqs, 388992284 residues

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	471.5	42.4	89	16 US-10-425-115-346968	Sequence 346968,
2	331	29.7	781	10 US-09-866-050A-668	Sequence 668, App
3	328	29.5	616	15 US-10-108-260A-3095	Sequence 3095, App
4	273	24.5	97	9 US-09-925-298-655	Sequence 655, App
5	273	24.5	97	14 US-10-102-806-655	Sequence 655, App
6	272	24.4	197	16 US-10-723-860-1946	Sequence 1946, App
7	222.5	20.0	511	15 US-10-104-047-2339	Sequence 2339, App
8	200	18.0	138	14 US-10-106-698-5959	Sequence 5959, App
9	188	16.9	83	9 US-09-864-761-44293	Sequence 44293, A
10	147	13.2	30	9 US-09-808-602-62	Sequence 62, Appl
11	123.5	11.1	826	15 US-10-395-241-18	Sequence 18, Appl

12	123	11.1	168	14	US-10-360-053-2	Sequence 2, Appl
13	123	11.1	168	14	US-10-360-053-4	Sequence 4, Appl
14	119	10.7	21	14	US-10-010-160-68	Sequence 68, Appl
15	119	10.7	23	17	US-10-495-715-70	Sequence 70, Appl
16	101	9.1	20	14	US-10-209-187A-18	Sequence 18, Appl
17	100.5	9.0	151	14	US-10-360-053-10	Sequence 10, Appl
18	100.5	9.0	159	14	US-10-360-053-12	Sequence 12, Appl
19	95	8.5	526	9	US-09-731-221-79	Sequence 79, Appl
20	93	8.4	289	17	US-10-688-745-10	Sequence 10, Appl
21	93	8.4	516	9	US-09-804-626-4	Sequence 4, Appl
22	93	8.4	519	15	US-10-359-369-34	Sequence 34, Appl
23	92	8.3	381	14	US-10-141-531-47	Sequence 47, Appl
24	92	8.3	381	14	US-10-141-531-48	Sequence 48, Appl
25	92	8.3	381	14	US-10-141-531-49	Sequence 49, Appl
26	92	8.3	381	14	US-10-141-531-50	Sequence 50, Appl
27	92	8.3	381	14	US-10-141-531-51	Sequence 51, Appl
28	92	8.3	381	14	US-10-141-531-52	Sequence 52, Appl
29	92	8.3	381	14	US-10-141-531-53	Sequence 53, Appl
30	92	8.3	381	14	US-10-141-531-54	Sequence 54, Appl
31	92	8.3	381	14	US-10-141-531-56	Sequence 56, Appl
32	92	8.3	381	14	US-10-141-531-58	Sequence 58, Appl
33	92	8.3	381	15	US-10-290-072-47	Sequence 47, Appl
34	92	8.3	381	15	US-10-290-072-48	Sequence 48, Appl
35	92	8.3	381	15	US-10-290-072-49	Sequence 49, Appl
36	92	8.3	381	15	US-10-290-072-50	Sequence 50, Appl
37	92	8.3	381	15	US-10-290-072-51	Sequence 51, Appl
38	92	8.3	381	15	US-10-290-072-52	Sequence 52, Appl
39	92	8.3	381	15	US-10-290-072-53	Sequence 53, Appl
40	92	8.3	381	15	US-10-290-072-54	Sequence 54, Appl
41	92	8.3	381	15	US-10-290-072-56	Sequence 56, Appl
42	92	8.3	381	15	US-10-290-072-58	Sequence 58, Appl
43	92	8.3	481	14	US-10-141-531-59	Sequence 59, Appl
44	92	8.3	481	15	US-10-290-072-59	Sequence 59, Appl
45	89.5	8.0	351	17	US-10-899-551-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1

US-10-425-115-346968
; Sequence 346968, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 346968
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_79599C.1.pep
US-10-425-115-346968

Query Match	42.4%	Score 471.5;	DB 16;	Length 89;
Best Local Similarity	90.0%;	Pred No. 7.7e-41;		
Matches	81;	Conservative	4;	Mismatches
				Indels
				Gaps
1	1	MGKTFYEDRGFGQHCYECSSDPCNLQYFRCNSIRVDSGCWMLYERPNYQGHQYFLRR	60	
1	1	MGKTFYEDRGFGQHCYECSSDPCNLQYFRCNSIRVDSGCWMLYERPNYQGHQYFLRR	60	
61	61	GDYDPYQWGMGFNSIRSCRLIPQHTGTFR	90	
61	61	GDYDPYQWGMGFNSIRSCRLIP-HTGSHR	89	

```
RESULT 2
US-09-866-050A-668
; Sequence 668, Application US/09866050A
; Publication No. US2003040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 668
; LENGTH: 781
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-668

Query Match      29.7%; Score 331; DB 10; Length 781;
Best Local Similarity 34.0%; Pred. No. 4.1e-25;
Matches 67; Conservative 33; Mismatches 77; Indels 20; Gaps 2;

QY 3 KITFYEDRGFQCHGYECSSDCPNLQ-----PYFSRCNSIRVDSGCWMLYERPNYQGHQ 55
Db 289 KAVVYAEFGFQSQSWEVSDIYNLQPEDSQSPQLTSVGLSLILGCGWGYEKEGFRGHQ 348

QY 56 YFLRGDYPDYQOWMGFNDISIRCLIPQHTGTFMRIRYERDDFRGQMSIITDDCPSLQD 115
Db 349 YLLEEGEYADWGHGDELLTSLRVIRTDGDPVAVLFPEDMDFQCHRVESVSSALPDVEL 408

QY 116 RHLTEVHSLNVLEGSWVLYEMPSYGRQYLLRPGEYRYLLDWGAMNAKVGSLRRV---- 171
Db 409 AQHGPSTQAIHLVSGVWVAYRGVFGSGEQYLEKGYRNCDDWGSNCALGSLQPVVQVG 468

QY 172 -----MDFYSDPN 179
Db 469 ESDLHFVTKIQLFSGPN 485

RESULT 3
US-10-108-260A-3095
; Sequence 3095, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3095
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3095

Query Match      29.5%; Score 328; DB 15; Length 616;
Best Local Similarity 33.8%; Pred. No. 6.2e-25;
Matches 70; Conservative 33; Mismatches 78; Indels 26; Gaps 2;

QY 7 YEDRFQCHGYECSSDCPNLQ-----PYFSRCNSIRVDSGCWMLYERPNYQGHQYFLR 59
Db 128 YEAPGFQGRSWEVSRDIYNLQPEDSQSPHLASVGLRVLGGCWVGYEKEGFRGHQYLL 187

QY 60 RGDYDPDYQOWMGFNDISIRCLIPQHTGTFMRIRYERDDFRGQMSIITDDCPSLQDRFHL 119
Db 188 EGEYDPDWSHGCGYDELLTSLRVIRTDGDPVAVLFPEDMDFEGHGVGVSKALPDVELVQHG 247

QY 120 TEVHSLNVLEGSWVLYEMPSYGRQYLLRPGEYRYLLDWGAMNAKVGSLRRVM----- 172
Db 248 PSTQAIHLVSGVWVAYQVGFSGEQYLEKGYRNCDDWGSNCALGSLQPVLVQGEHDL 307

QY 173 -----DFYSDPNSSSVKLA 187
Db 308 HFSKIQLFSPDFGLGDHFSFEDDQAA 334

RESULT 4
US-09-925-298-655
; Sequence 655, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 655
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (91)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-655

Query Match      24.5%; Score 273; DB 9; Length 97;
Best Local Similarity 53.8%; Pred. No. 3e-20;
Matches 43; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 3 KITFYEDRGFQCHGYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRRGD 62
Db 15 QITFYEDKNFGQRRYDCDCDCADKHXYLSRCNSIKVEGTTWAVYERPFPAGVMTLPQGE 74

QY 63 YPDYQOWMGFNDISIRCLIP 82
Db 75 YPEYQRMWGLNRLSLXRAV 94

RESULT 5
US-10-102-806-655
; Sequence 655, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 655
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (91)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-655

Query Match          24.5%; Score 273; DB 14; Length 97;
Best Local Similarity 53.8%; Pred. No. 3e-20;
Matches 43; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY   3 KITFYEDRGFGCHGYECSSDCPNLQPYFSCNSIRVDSGCWMLYERPNYQHGYFLRRGD 62
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db    15 QITFYEDKNFGRRYDCCDACDHLYLSCNSIKVEGGTWAVERPFAFYMYILPQQE 74
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   63 YPDYQQWMGFNDIRSICRLI 82
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db    75 YPEYQRWGLNDRLLSXRAV 94
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

RESULT 6
US-10-723-860-1946
; Sequence 1946, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1946
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-1946

Query Match          24.4%; Score 272; DB 16; Length 197;
Best Local Similarity 34.4%; Pred. No. 9.3e-20;
Matches 63; Conservative 38; Mismatches 66; Indels 16; Gaps 7;

QY   4 ITFYEDRGFGCHGYECSSDCPNL--QPYFSRCNSIRVDSCWMLYERPNYQHGYFLRRG 61
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db    14 LTLDEEDFOGRRCLLLSDCANVCERGGLPRVRSVKVENGVWVAFEYPDFGQQQFILEKG 73
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   62 DYPDYQQWMGF----NDSIRSCR--LIPOHTGTFRMRIVERDDPRGOMSEITDDCPSLOD 115
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db    74 DYPRSAWSGSSSHNGSLSPRVLCANHANDS-RVTLPFGDNFQGCKFDLVDDYPSLPS 132
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   116 R-PHLTEVLSNLVLGESSWLVIEMPSYRGROYLF---RPEYRYLDWG--ANNKVGSGL 168
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db    133 MGWASKDVGLSKVSSGANWAYQYPGYRGYYQLVLRDRHSGEFTCYELGTQAHTGQLQSI 192
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY   169 RVV 171
      ||||
Db    193 RVV 195
      ||||

RESULT 7
US-10-104-047-2339
; Sequence 2339, Application US/10104047
```



```
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: PET-28a-c(+) cloning/expression region
US-10-495-715-70
```

```
Query Match      10.7%; Score 119; DB 17; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      177 DNSSSVDKLAAALEHHHHH 197
          |||||
Db       3 DNSSSVDKLAAALEHHHHH 23
```

```
Search completed: July 25, 2005, 17:00:47
Job time : 90.8485 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:05:40 ; Search time 101.007 Seconds
(without alignments)
758.153 Million cell updates/sec

Title: US-10-030-605A-19
Perfect score: 1105
Sequence: 1 AAQPMGRKPKEDRGFGH.....FYAAGAPVYPDLEPRAA 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_16Dec04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	198	4	AAB46848
2	1063	96.2	198	4	AAB46849
3	986	89.2	197	4	AAB46850
4	944	85.4	197	4	AAB46851
5	769.5	69.6	173	7	ADD47108
6	769.5	69.6	173	7	AD57551
7	769.5	69.6	173	7	ADD47112
8	769.5	69.6	173	7	ADD47112
9	769.5	69.6	173	7	ADD47118
10	769.5	69.6	173	7	AD583335
11	769.5	69.6	173	7	AD57555
12	757.5	68.6	173	7	ADD47104
13	727.5	65.8	174	4	ABG21006
14	727.5	65.8	174	7	ADD47106
15	714.5	64.7	362	7	ADC31213
16	710.5	64.3	173	7	ADD47110
17	710.5	64.3	173	7	AD583337
18	710.5	64.3	173	7	ADD47116
19	710.5	64.3	173	7	AD57553
20	705	63.8	308	4	ABG21004
21	682	61.7	199	4	ABG21007
22	604	54.7	145	4	ABG21005
23	335	30.3	169	8	ADQ97976
24	321.5	29.1	507	8	ADC31503
25	321.5	29.1	616	7	ADM04410

26	321.5	29.1	888	8	ADO44168	Ado44168	Structura
27	317.5	28.7	781	5	ABB72344	Abb72344	Murine pr
28	314	28.4	420	5	ABP69298	Abp69298	Human pol
29	304	27.5	204	7	AD663507	Ad663507	Rat Prote
30	304	27.5	205	8	ADP22671	Adp22671	Golden ha
31	300	27.1	205	2	AAW17523	Aaw17523	Human bet
32	297	26.9	211	8	ADP22673	Adp22673	Golden ha
33	292	26.4	252	5	AAU11447	Aau11447	Human cry
34	281.5	25.5	215	2	AAW17522	Aaw17522	Human bet
35	280	25.3	250	8	ADP22669	Adp22669	Golden ha
36	273	24.7	542	6	ABR41645	Abr41645	Human DIT
37	273	24.7	962	7	AD229065	Ad229065	Human AIT
38	273	24.7	1723	6	ABG74680	Abg74680	Human CGD
39	271	24.5	1080	4	AAW00803	Aaw00803	Human bon
40	271	24.5	1637	4	AAW00916	Aaw00916	Human bon
41	271	24.5	2263	4	AAW79000	Aaw79000	Human pro
42	266	24.1	197	4	AAW47135	Aaw47135	CDIFF-16,
43	264	23.9	197	4	ABG08246	Abg08246	Novel hum
44	264	23.9	197	8	ADQ19127	Adq19127	Human sof
45	253	22.9	97	3	AAB58947	Aab58947	Breast an

ALIGNMENTS

RESULT 1
AAB46848
ID AAB46848 standard; protein; 198 AA.
XX
AC AAB46848;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline mutant Mu12A protein fragment.
XX
XX Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW Biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus
XX
PW DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
RA (FJED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
DR WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 16-17; 28pp; German.
XX
CC This invention describes a novel protein (I) with beta-'leaflet',
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene

CC therapy vectors and for intracellular immunization. (I) can be provided
 CC with new or improved specific antibody-like binding, catalytic or
 CC fluorescent properties, without the cost and difficulties associated with
 CC producing complete or recombinant antibodies. (I) are relatively small
 CC (20 kDa) and can be expressed with other components as multifunctional
 CC fusions. They have good stability against low pH, denaturing agents and
 CC high temperatures, conditions under which antibodies are unstable
 XX
 SQ Sequence 198 AA;

Query Match 100.0%; Score 1105; DB 4; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.1e-110;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAQPMAGRIKFEDRGFGQGHYSCNSDCPNLQPFYFRCNSIRVLSGCWMLYERPNYQGHQ 60
 DB 1 AAQPMAGRIKFEDRGFGQGHYSCNSDCPNLQPFYFRCNSIRVLSGCWMLYERPNYQGHQ 60
 QY 61 YFLRRGDYDPYQOWMGFNDISIRSCRLIPOHTGTFMRRIYERDDFRGQSEITDDCPSLQD 120
 DB 61 YFLRRGDYDPYQOWMGFNDISIRSCRLIPOHTGTFMRRIYERDDFRGQSEITDDCPSLQD 120
 QY 121 RFHLTEVHSLNVLEGSWLVYEMPSYRGROYLLRPGYRRLYLDWGAMNAKVGSLRRVMDFY 180
 DB 121 RFHLTEVHSLNVLEGSWLVYEMPSYRGROYLLRPGYRRLYLDWGAMNAKVGSLRRVMDFY 180
 QY 181 AAAGAPVPYDPLEPRAA 198
 DB 181 AAAGAPVPYDPLEPRAA 198

RESULT 2
 AAB46849
 ID AAB46849 standard; protein; 198 AA.
 XX
 AC AAB46849;
 XX
 DT 26-APR-2001 (first entry)
 XX
 DE Bovine gamma-crystalline protein fragment.
 XX
 KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
 KW biosensor; pollution detection; pollution control; gene therapy;
 KW intracellular immunization.
 XX
 OS Bos taurus.

DE19932688-A1.
 18-JAN-2001.
 PF 13-JUL-1999; 99DE-01032688.
 PR 13-JUL-1999; 99DE-01032688.
 PA (FIED//) FIEDLER U.
 PA (RUDO//) RUDOLPH R.
 XX
 PI Rudolph R, Fiedler U, Boehm G, Reimann C;
 XX
 DR WPI; 2001-148304/16.
 XX
 PT Mutated proteins having beta-leaflet structure and related nucleic acid,
 PT have new or improved properties, e.g. antibody-like specific binding or
 PT catalytic activity.
 XX
 PS Claim 11; Page 17-18; 28pp; German.

CC This invention describes a novel protein (I) with beta-'leaflet'
 CC structure having surface-exposed amino acids, present in at least two
 CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
 CC protein is altered by targeted mutagenesis so that it has new, or
 CC improved, specific binding, catalytic or fluorescent properties. The

CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
 CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
 CC contain (II) or (III), or their fragments that encode a functional region
 CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
 CC and therapy, in cosmetics, bioseparation and biosensors, and for
 CC pollution detection and control, e.g. for specific targeting of gene
 CC therapy vectors and for intracellular immunization. (I) can be provided
 CC with new or improved specific antibody-like binding, catalytic or
 CC fluorescent properties, without the cost and difficulties associated with
 CC producing complete or recombinant antibodies. (I) are relatively small
 CC (20 kDa) and can be expressed with other components as multifunctional
 CC fusions. They have good stability against low pH, denaturing agents and
 CC high temperatures, conditions under which antibodies are unstable
 XX

SQ Sequence 198 AA;

Query Match 96.2%; Score 1063; DB 4; Length 198;
 Best Local Similarity 96.5%; Pred. No. 3.8e-106;
 Matches 191; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AAQPMAGRIKFEDRGFGQGHYSCNSDCPNLQPFYFRCNSIRVLSGCWMLYERPNYQGHQ 60
 DB 1 AAQPMAGRIKFEDRGFGQGHYSCNSDCPNLQPFYFRCNSIRVLSGCWMLYERPNYQGHQ 60
 QY 61 YFLRRGDYDPYQOWMGFNDISIRSCRLIPOHTGTFMRRIYERDDFRGQSEITDDCPSLQD 120
 DB 61 YFLRRGDYDPYQOWMGFNDISIRSCRLIPOHTGTFMRRIYERDDFRGQSEITDDCPSLQD 120
 QY 121 RFHLTEVHSLNVLEGSWLVYEMPSYRGROYLLRPGYRRLYLDWGAMNAKVGSLRRVMDFY 180
 DB 121 RFHLTEVHSLNVLEGSWLVYEMPSYRGROYLLRPGYRRLYLDWGAMNAKVGSLRRVMDFY 180
 QY 181 AAAGAPVPYDPLEPRAA 198
 DB 181 AAAGAPVPYDPLEPRAA 198

RESULT 3
 AAB46850
 ID AAB46850 standard; protein; 197 AA.
 XX
 AC AAB46850;
 XX
 DT 26-APR-2001 (first entry)
 XX
 DE Bovine gamma-crystalline mutant protein Mu12A-His fragment.
 XX
 KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
 KW biosensor; pollution detection; pollution control; gene therapy;
 KW intracellular immunization.
 XX
 OS Bos taurus.

DE19932688-A1.
 18-JAN-2001.
 PF 13-JUL-1999; 99DE-01032688.
 PR 13-JUL-1999; 99DE-01032688.
 PA (FIED//) FIEDLER U.
 PA (RUDO//) RUDOLPH R.
 XX
 PI Rudolph R, Fiedler U, Boehm G, Reimann C;
 XX
 DR WPI; 2001-148304/16.

PT Mutated proteins having beta-leaflet structure and related nucleic acid,
 PT have new or improved properties, e.g. antibody-like specific binding or
 PT catalytic activity.
 XX
 PS Claim 11; Page 18-19; 28pp; German.

XX This invention describes a novel protein (I) with beta-'leaflet'
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene
CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;

Query Match 89.2%; Score 986; DB 4; Length 197;
Best Local Similarity 99.4%; Pred. No. 7.7e-98;
Matches 175; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 6 MGRIFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYERPNYQGHQYFLRR 65
DB 1 MGRIFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYERPNYQGHQYFLRR 60
QY 66 GDYPDYQQWGMGFNDISRCLIPQHTGTFMRMIYERDDFRGQMSITDDCPSLQDRFHULT 125
DB 61 GDYPDYQQWGMGFNDISRCLIPQHTGTFMRMIYERDDFRGQMSITDDCPSLQDRFHULT 120
QY 126 EVHSLNVLGSGWLVYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 181
DB 121 EVHSLNVLGSGWLVYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 176

RESULT 4
AAB46851
ID AAB46851 standard; protein; 197 AA.
XX
AC AAB46851;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline protein WT-HIS fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
XX Rudolph R, Fiedler U, Boehm G, Reimann C;
PI WPI; 2001-148304/16.
XX
DR Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.

XX Claim 11; Page 19-20; 28pp; German.
XX
CC This invention describes a novel protein (I) with beta-'leaflet',
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene
CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;
Query Match 85.4%; Score 944; DB 4; Length 197;
Best Local Similarity 95.5%; Pred. No. 2.6e-93;
Matches 168; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 6 MGRIFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYERPNYQGHQYFLRR 65
DB 1 MGRIFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYERPNYQGHQYFLRR 60
QY 66 GDYPDYQQWGMGFNDISRCLIPQHTGTFMRMIYERDDFRGQMSITDDCPSLQDRFHULT 125
DB 61 GDYPDYQQWGMGFNDISRCLIPQHTGTFMRMIYERDDFRGQMSITDDCPSLQDRFHULT 120
QY 126 EVHSLNVLGSGWLVYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 181
DB 121 EVHSLNVLGSGWLVYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFYA 176
RESULT 5
ADD47108
ID ADD47108 standard; protein; 173 AA.
XX
AC ADD47108;
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P10065, SEQ ID NO 12798.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
OS Unidentified.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX

```
DR WPI; 2003-268312/26.
DR GENBANK; P10065.
XX
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX
XX PS Example 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 173 AA;

Query Match          69.6%; Score 769.5; DB 7; Length 173;
Best Local Similarity 77.6%; Pred. No. 1.5e-74;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFEDRGFGQHYHSCNSDCPNLQYFSCRNSIRVLSGCWMLYERPNYQGHQYFLRRG 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GKITFEDRGFGQRCYECSCDCPNLQYFSCRNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

QY 67 DYPDYQQWNGFNDISRCLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRPHLTE 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 DYPDYQQWNGFSDSIRSCRSIP-YTSSHRILRYERDDYRGLVSELTCSCIHDRFLRNE 119

QY 127 VHSLNVLGSGWLYEMPSYRGROYLLRPGYRRLYLDWGAMNAKVGLSRVMDIFY 180
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 IYSHMVLEGSWLYEMPNYRGQYLLRPGDYRRYHDWGAMDAKVGLSRVMDLY 173

RESULT 6
AD857551
ID ADE57551 standard; protein; 173 AA.
AC ADE57551;
XX
XX 29-JAN-2004 (first entry)
XX
XX Rat Protein P10065, SEQ ID NO 3413.
DE
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
```

```
XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX
XX PA (GHO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX DR WPI; 2003-268312/26.
XX DR GENBANK; P10065.
XX
XX PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 173 AA;

Query Match          69.6%; Score 769.5; DB 7; Length 173;
Best Local Similarity 77.6%; Pred. No. 1.5e-74;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFEDRGFGQHYHSCNSDCPNLQYFSCRNSIRVLSGCWMLYERPNYQGHQYFLRRG 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GKITFEDRGFGQRCYECSCDCPNLQYFSCRNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

QY 67 DYPDYQQWNGFNDISRCLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRPHLTE 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 DYPDYQQWNGFSDSIRSCRSIP-YTSSHRILRYERDDYRGLVSELTCSCIHDRFLRNE 119

QY 127 VHSLNVLGSGWLYEMPSYRGROYLLRPGYRRLYLDWGAMNAKVGLSRVMDIFY 180
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 IYSHMVLEGSWLYEMPNYRGQYLLRPGDYRRYHDWGAMDAKVGLSRVMDLY 173

RESULT 7
ADD47114
ID ADD47114 standard; protein; 173 AA.
XX
XX AC ADD47114;
XX
```


CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 173 AA;

Query Match 69.6%; Score 769.5; DB 7; Length 173;
Best Local Similarity 77.6%; Pred. No. 1.5e-74;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFEDRGFGQHYYSNCSDCPNLQYFSCNSIRVLSGCWMLYERNYQGHQYFLRRG 66

Db 1 GKITFYEDRGFGQRCYECSSDCPNLQYFSCNSIRVDSGCWMLYERNYQGYFLRRG 60

QY 67 DYPDYQWGMGFNDISIRSCRLIPQHTGTFRMIYERDDFRGQMSITDDCPSLQDRFHLTE 126

Db 61 DYPDYQWGMGFSDSIRSCRSIP-VTSSHRIRLYERDDYRGLVSELTDCSCIHDRFRLNE 119

QY 127 VHSNLVLEGSWLVYEMPSYRGQYLLRPGCYRRYLDWGAMNAKVGSLRRVMDFY 180

Db 120 IYSMHVLEGSWLVYEMPYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173

RESULT 9

ADD47118

ID ADD47118 standard; protein; 173 AA.

XX AC ADD47118;

XX XX

DT 02-DEC-2004 (revised)

DT 29-JAN-2004 (first entry)

XX DE

DE Rat Protein AAA40981, SEQ ID NO 12808.

XX XX

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS

OS Rattus norvegicus.

OS Unidentified.

XX XX

XX WO2003016475-A2.

XX XX

XX 27-FEB-2003.

XX XX

XX 14-AUG-2002; 2002WO-US025765.

XX XX

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX XX

XX (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX XX

XX Woolf C, D'urso D, Befort K, Costigan M;

XX XX

XX WPI: 2003-268312/26.

DR GENBANK; AAA40981.

XX XX

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX XX

XX Example 1; Page; 1017pp; English.

XX XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 173 AA;

Query Match 69.6%; Score 769.5; DB 7; Length 173;

Best Local Similarity 77.6%; Pred. No. 1.5e-74;

Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFEDRGFGQHYYSNCSDCPNLQYFSCNSIRVLSGCWMLYERNYQGHQYFLRRG 66

Db 1 GKITFYEDRGFGQRCYECSSDCPNLQYFSCNSIRVDSGCWMLYERNYQGYFLRRG 60

QY 67 DYPDYQWGMGFNDISIRSCRLIPQHTGTFRMIYERDDFRGQMSITDDCPSLQDRFHLTE 126

Db 61 DYPDYQWGMGFSDSIRSCRSIP-VTSSHRIRLYERDDYRGLVSELTDCSCIHDRFRLNE 119

QY 127 VHSNLVLEGSWLVYEMPSYRGQYLLRPGCYRRYLDWGAMNAKVGSLRRVMDFY 180

Db 120 IYSMHVLEGSWLVYEMPYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173

RESULT 10

ADE83335

ID ADE83335 standard; protein; 173 AA.

XX AC ADE83335;

XX XX

XX 29-JAN-2004 (first entry)

XX DT

XX Rat Protein P10065, SEQ ID NO 10928.

XX DE

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX KW

XX Rattus norvegicus.

XX OS

XX WO2003016475-A2.

XX XX

XX 27-FEB-2003.

XX XX

XX 14-AUG-2002; 2002WO-US025765.

XX XX

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX XX

XX (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX XX

XX Woolf C, D'urso D, Befort K, Costigan M;

XX XX

XX WPI: 2003-268312/26.

DR GENBANK; P10065.

XX XX

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX XX

XX Claim 1; Page; 1017pp; English.

Db 61 GDYADHQWGLSDSVSRCLIP-HSGSHRIRLYEREDYRGQMIETEDCSCLQDRFRN 119

QY 126 EVHSLNVLGSGWLYEMPYRGQYLLRPGYRRYLDWGAMNAKVGSLRVNMF 179

Db 120 EIHSLNVLGSGWLYELSNYRGQYLLMPGDYRRYQDWGATNARVGSLLRRVIDF 173

RESULT 14

ID ADD47106 standard; protein; 174 AA.

AC ADD47106;

DT 02-DEC-2004 (revised)

DT 29-JAN-2004 (first entry)

XX Human Protein NP_008822, SEQ ID NO 12796.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

OS Unidentified.

XX WO2003016475-A2.

PN 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI: 2003-268312/26.

DR GENBANK; NP_008822.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the

CC activity in an animal of one or more of the polypeptides given in the

CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more

CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

CC therapy). The sequence presented is a human protein (described in table 3

CC of the specification) which is differentially expressed during pain.

CC Note: The sequence data for this patent did not form part of the printed

CC

CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 174 AA;

Query Match 65.8%; Score 727.5; DB 7; Length 174;

Best Local Similarity 73.6%; Pred. No. 5.3e-70;

Matches 128; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 6 MGRIFKEDRGFGQGHYYSCNSDCPNLQPYFSRCNSIRVLGCMWMLYERPNYGHGYELRR 65

Db 1 MKITLYEDRGFGQGHYYSCNSDCPNLQPYFSRCNSIRVLGCMWMLYERPNYGHGYELRR 60

QY 66 GDYDYQWGMGFNDISIRSCRLIPQHTGTFMRIRYERDDFRGQMSITDDCPSLQDRFHLT 125

Db 61 GDYADHQWGLSDSVSRCLIP-HSGSHRIRLYEREDYRGQMIETEDCSCLQDRFRN 119

QY 126 EVHSLNVLGSGWLYEMPYRGQYLLRPGYRRYLDWGAMNAKVGSLRVNMF 179

Db 120 EIHSLNVLGSGWLYELSNYRGQYLLMPGDYRRYQDWGATNARVGSLLRRVIDF 173

RESULT 15

ID ADC31213

XX ADC31213 standard; protein; 362 AA.

AC ADC31213;

DT 18-DEC-2003 (first entry)

XX Human novel polypeptide sequence, SEQ ID NO:1295.

XX Human; diagnostic; drug screening; forensic; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy; chromosome 18.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX Tang TV, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

PI Haley-Vicente D, Drmanac RT;

XX WPI: 2003-371981/35.

DR N-FSDB; ADC30242.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or

PT treating conditions such as neurodegenerative diseases, anemias, platelet

PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

PT cancer.

XX Claim 20; SEQ ID NO 1295; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-

CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The

CC invention also relates to nucleic acid sequences over 99% identical with

CC the novel human cDNAs. The invention additionally encompasses expression

CC vectors and host cells comprising a nucleic acid of the invention; the

CC recombinant production of a polypeptide of the invention; an antibody

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:18:55 ; Search time 26.0878 Seconds
(without alignments)
566.567 Million cell updates/sec

Title: US-10-030-605A-19
Perfect score: 1105
Sequence: 1 AAQPAWGRKIKEDRGFGQH.....FYAAGAPVVPDPLEPRAA 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep:**
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep:**
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep:**
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep:**
- 5: /cgn2_6/prodata/1/iaa/ECTUS COMB.pep:**
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830	75.1	187	4	US-09-949-016-9254
2	747.5	67.6	179	4	Sequence 9254, Ap
3	532.5	48.2	192	4	Sequence 11310, A
4	300	27.1	205	2	US-08-729-152-8
5	296	26.8	209	4	Sequence 8, Appli
6	294.5	26.7	204	4	Sequence 10830, A
7	292	26.4	262	4	Sequence 9696, Ap
8	290.5	26.3	207	4	Sequence 10815, A
9	287	26.0	235	4	Sequence 10814, A
10	281.5	25.5	215	2	US-08-729-152-1
11	96.5	8.7	257	4	Sequence 1, Appli
12	96.5	8.7	257	4	Sequence 2, Appli
13	96.5	8.7	297	4	Sequence 2, Appli
14	96	8.7	18	4	Sequence 12, Appli
15	96	8.7	43	2	US-08-486-814A-2
16	96	8.7	256	4	Sequence 43, Appl
17	96	8.7	258	4	Sequence 2, Appli
18	79.5	7.2	139	4	Sequence 4, Appli
19	78.5	7.1	243	4	Sequence 5088, Ap
20	78.5	7.1	516	4	Sequence 7356, Ap
21	78	7.1	48	6	Sequence 10341, A
22	78	7.1	48	6	Patent No. 5168041-1
23	78	7.1	49	6	Patent No. 5168041
24	78	7.1	49	6	Patent No. 5164483-2
25	78	7.1	49	6	Patent No. 5434245-1
26	78	7.1	49	6	Patent No. 5164483-2
27	78	7.1	50	6	Patent No. 5434245-1

28	78	7.1	50	6	5434245-3	Patent No. 5434245
29	78	7.1	50	6	5434245-2	Patent No. 5434245
30	78	7.1	50	6	5434245-3	Patent No. 5434245
31	78	7.1	51	6	5434245-4	Patent No. 5434245
32	78	7.1	51	6	5434245-4	Patent No. 5434245
33	76	6.9	13	2	US-08-796-598-13	Sequence 13, Appl
34	76	6.9	13	2	US-08-447-175A-13	Sequence 13, Appl
35	76	6.9	13	2	US-08-943-915-6	Sequence 6, Appli
36	76	6.9	13	3	US-08-881-037-112	Sequence 112, App
37	76	6.9	13	4	US-09-693-945-14	Sequence 14, Appl
38	73	6.6	7831	4	US-09-902-540-12902	Sequence 12902, A
39	72.5	6.6	436	2	US-08-899-028A-2	Sequence 2, Appli
40	72.5	6.6	436	2	US-09-210-124-2	Sequence 2, Appli
41	72.5	6.6	436	3	US-09-210-009-2	Sequence 2, Appli
42	72	6.5	49	6	5164483-1	Patent No. 516448
43	72	6.5	49	6	5164483-1	Patent No. 516448
44	72	6.5	98	6	5164483-3	Patent No. 516448
45	72	6.5	98	6	5164483-3	Patent No. 516448

ALIGNMENTS

RESULT 1
US-09-949-016-9254
; Sequence 9254, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9254
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9254

Query Match	75.1%	Score 830;	DB 4;	Length 187;
Best Local Similarity	81.7%	Pred. No. 6.6e-82;		
Matches 143;	Conservative 16;	Mismatches 16;	Indels 0;	Gaps 0;
QY	6	MCRIKFKEDRGFGQHYHSCNDCPNLQYFSCNIRVLSCGCMWLYERPNYQGHQYFLRR	65	
DB	13	MKGIIFYEDRAFQGRSYECTTDCPNLQYFSCNIRVESGCMWLYERPNYQGHQYFLRR	72	
QY	66	GYDPYQQWGMGNDIRSCLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRFHLT	125	
DB	73	GEYDPYQQWGLSUDSIRSCLIPPHSGAYRMKIYDRDELRCQMSLTDCLSVQDRFHLT	132	
QY	126	EVHSLNVLEGSVWLYEMPYSYRGRQYLLRPGYRRYLLDGMAMNAKVGSLLRRVMDFY	180	
DB	133	EIHSNLVLEGSWLYEMPYRGRQYLLRPGYRRYLLDGMAMNAKVGSLLRRVMDLY	187	

RESULT 2
US-09-949-016-11310
; Sequence 11310, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11340
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11340

Query Match          67.6%; Score 747.5; DB 4; Length 179;
Best Local Similarity 74.2%; Pred. No. 5.3e-73;
Matches 132; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 2 AQPAMGRIKEDRGFGQHHYSCNDCPNLQPYFGRCSIRVLSGCMWLYERPNTQGHQY 61
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 AQPAMGKITLYEDRGFGQHHYSCSDHPNLQPYLSRCSARVDSGCMWLYEQPNYSLQY 61
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 62 FLRRGDYPPYQWGMFENDSIRSCRLIPQHTGTFMRRIYERDDFRGOMSEITDDCPSLQDR 121
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 62 FLRRGDYADHQWGLSDSVSRCLIP-HSGSHRIRLRYERDYRGOMTEFTEDCSLQDR 120
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 122 FHLTEVHLSNVLEGSWLVYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLRRVMD 179
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 FRFNEIHSNVLEGSWLVYELNSYRGROYLLRPGYRRYLDWGATNARVGSLLRVDF 178
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 3
US-09-949-016-11340
; Sequence 11340, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11340
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11340

Query Match          48.2%; Score 532.5; DB 4; Length 192;
Best Local Similarity 51.2%; Pred. No. 1.1e-49;
Matches 88; Conservative 36; Mismatches 47; Indels 1; Gaps 1;

QY 8 RIKFKEDRGFGQHHYSCNDCPNLQPYFGRCSIRVLSGCMWLYERPNTQGHQYFLRRGD 67
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 21 KITFYEDKNFQGRRYDCDCADFTYLSRCSIKVEGGTWAVERPFAFYMYILPQGE 80
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 68 YPDYQWGMFENDSIRSCRLIPQHT-GTFMRRIYERDDFRGOMSEITDDCPSLQDRFHLTE 126
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 81 YPEYQWGMGLNRLSSCRAVHLPSGGQYKIQIFKGFQGMVYETTEDCPSEMQPHMRE 140
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 127 VHSLSNVLEGSWLVYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLRRVMD 178
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
Db 141 IHSCKVLGVWIFYELPNYRGROYLLDKCKYRKPIDWGAASPAVQSFRRIVE 192

RESULT 4
US-08-729-152-8
; Sequence 8, Application US/08729152
; Patent No. 5871739
; GENERAL INFORMATION:
; APPLICANT: Inoue, Eri
; TITLE OF INVENTION: Pharmaceutical Composition
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: Southern Building, Suite 700, 805 Fifteenth
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,152
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: JP 291993/1995
; FILING DATE: 13-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
US-08-729-152-8

Query Match          27.1%; Score 300; DB 2; Length 205;
Best Local Similarity 37.1%; Pred. No. 1.8e-24;
Matches 63; Conservative 31; Mismatches 72; Indels 4; Gaps 3;

QY 13 EDRGFGQHHYSCNDCPNL-QPYFSRCNSIRVLSGCMWLYERPNTQGHQYFLRRGDYDPY 71
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 23 EQENFGHSHELNGPCPNLKETGVEKAGSVLVQAGPWVGVEQANCKGQGFVEKEGYPW 82
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 72 QOWMGF--NDSIRSCRLIPQHTGTFMRRIYERDDFRGOMSE-ITDDCPSLQDRFHLTEVH 128
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 83 DSWTSSRRSTDLSLSLRPIKVDQSQEHKILYENPNFTGKMEIIDDDVPSFHAHGYQEKVS 142
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 129 SLNVLEGSWLVYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLRRVMD 178
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 143 SVRVQSGTWVGYPGYRGLQLLEKDYKOSDSFGAPHPQVQSVRRIRD 192
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 5
US-09-949-016-10830
; Sequence 10830, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10830
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10830

Query Match          26.8%; Score 296; DB 4; Length 209;
Best Local Similarity 37.1%; Pred. No. 5.1e-24;
Matches 63; Conservative 31; Mismatches 72; Indels 4; Gaps 3;

QY 13 EDGFGGHVYSCNDCPNL-QPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRRGDYDPY 71
Db 27 EQENFQGHSHLNGPCPNLKETGVKAGSVLVQAGPVGWYEQANCKGEQVFPERGEYPRW 86
QY 72 QQWMGF--NDSIRSCRLIPQHTGTFMRIRYERDDFRGQMSB-ITDDCPSLQDRFHLTEVH 128
Db 87 DSWTSRRRDSLSLRPIKVDQEHKILLYENPNTGKQWEIILDDVPSFHAHGQEKVS 146
QY 129 SUNLVGWSWLYEMPSYGRQYLLRPGYRRYLDWGAMNAKVGSLLRRVMD 178
Db 147 SVMVRSGTWVGYPGYRGLQVLLKGDYKDSDFGAPHPQVQSVRRIRD 196

RESULT 6
US-09-949-016-9696
; Sequence 9696, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9696
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9696

Query Match          26.7%; Score 294.5; DB 4; Length 204;
Best Local Similarity 35.5%; Pred. No. 7.2e-24;
Matches 65; Conservative 37; Mismatches 66; Indels 15; Gaps 7;

QY 8 RIKFKEDRGFGQHYHSCNDCPNL-QPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRRG 66
Db 21 KITIVDQENFQGRKEFTSCPNVSRSDNVRSLKVESGAWIGYHTSFQCGQQFILRG 80
QY 67 DYPDYQQWNGEN----DSIRSCRLI--POHTGTFMRIRYERDDFRGQMSB-ITDDCPSLQD 120
Db 81 EFPWRDAGSGSNAYHIERLMSFRPICSANHKES-KMTIFENKFNFIQRQWEISDDIPSLQA 139
QY 121 R-FHLTEVHSLNVLGWSWLYEMPSYGRQYLL----RPGYRRYLDWG--AMNAKVGSLL 173
```

```

Db 140 MGWFNFVSGMKIOSGAWVCYQPGYRGYQILECDHHRGDDYKHWREWSHAQTSIQSI 199
QY 174 RRV 176
Db 200 RRI 202

RESULT 7
US-09-949-016-10815
; Sequence 10815, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10815
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10815

Query Match          26.4%; Score 292; DB 4; Length 262;
Best Local Similarity 34.4%; Pred. No. 1.9e-23;
Matches 67; Conservative 34; Mismatches 88; Indels 6; Gaps 4;

QY 4 PAMGRIKFKEDRGFGQHYHSCNDCPNLQPY-FSRCNSIRVLSCWMLYERPNYQGHQYF 62
Db 66 PGNYRLVVFELNFQGRRAEFSGECNSLADRGDFRVSIIVSAGPWVAFQSNFRGEMFI 125
QY 63 LRGGDYDYPQWMGF--NDSIRSCRLIPQHTGTFMRIRYERDDFRGQMSB-ITDDCPSLQ 119
Db 126 LEKGEYPRWNTWSSYRSRDLMSFRPIKWDQAEHKISLFEKANFKGNTIEIQGDDAPSLW 185
QY 120 DRFHLTEVHSLNVLGWSWLYEMPSYGRQYLLRPGYRRYLDWGAMNAKVGSLLRRVMD- 178
Db 186 VYGFSDRVGVSQVSSGTWVGYPGYRGLQVLLKGDYKDSDFGAPHPQVQSVRRIRD 245
QY 179 -FYAAAGAPVPYPDP 192
Db 246 QWHLEGSFPVLATEP 260

RESULT 8
US-09-949-016-10814
; Sequence 10814, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 10814
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10814

Query Match          26.3%; Score 290.5; DB 4; Length 207;
Best Local Similarity 36.3%; Pred. No. 2e-23;
Matches 65; Conservative 37; Mismatches 60; Indels 17; Gaps 7;

QY 13 EDGFGQGHYSCNSDCPNL-LQPYFSCNSIRVLSCGMLYERPNTYQGHQYFLRRGDYDPY 71
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 29 DEDGFGRRHEFTAECPSSLGELGFTVRSLSKLVSGAWGFEHAGFGQGGQYILERGEYPSW 88
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY 72 QQWGFNDISRCRL-----IPQHTGTFRMRIYERDDFRGQMSHITDDCPSLQDR-FH 123
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 89 DAW-GGNTAIPAERLTSPFAACANHRDS-RLTIFEQENFLGKGELSDDYPSLQAMGWE 146
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY 124 LTEVHSLNVLEGSWLYEMPSYRGROYLL----RGEYRRLDWG--AMNAKVGSLRRV 176
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 147 GNEVGSFHVHSGAWVCSQFGIRGFGQYVLECDHHSDYKHFRWGSHAPTFQVQSIRRI 205
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

RESULT 9
US-09-949-016-8099
; Sequence 8099, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8099
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8099

Query Match          26.0%; Score 287; DB 4; Length 235;
Best Local Similarity 36.0%; Pred. No. 5.7e-23;
Matches 63; Conservative 32; Mismatches 66; Indels 14; Gaps 5;

QY 13 EDGFGQGHYSCNSDCPNL-QPYFSCNSIRVLSCGMLYERPNTYQGHQYFLRRGDYDPY 71
   |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 54 ELENFGKRCELSAECPSLTDSLLEKVGSIQVESGFWLAFESRAFGEQFVLEKGDYPRW 113
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY 72 QQWGF--NDSRSCRLIPQHTGTFRMRIYERDDFRGQMSHIT-DDCPSL-----QDRPH 123
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 114 DAWNSRSDSLSLRLNIDSPDKHLFENPAFSGRKXMEIVDDVPSLWAHGFQDR-- 171
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY 124 LTEVHSLNVLEGSWLYEMPSYRGROYLLRCPGEYRRLDWGAMNAKVGSLRRVMD 178
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 172 ---VASVRAINGTWGFEFGYRGROYVFERGEYRHNENWDASQQLQSVRRIRD 223
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

RESULT 10
US-08-729-152-1
; Sequence 1, Application US/08729152
; Patent No. 5871739
; GENERAL INFORMATION:
; APPLICANT: Inoue, Eri
; TITLE OF INVENTION: Pharmaceutical Composition
; NUMBER OF SEQUENCES: 47
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: Southern Building, Suite 700, 805 Fifteenth
; STREET: Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,152
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 291993/1995
; FILING DATE: 13-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
US-08-729-152-1

Query Match          25.5%; Score 281.5; DB 2; Length 215;
Best Local Similarity 34.1%; Pred. No. 2e-22;
Matches 62; Conservative 37; Mismatches 70; Indels 13; Gaps 6;

QY 8 RIKPKEDRGFGHYVSCNSDCPNL-QPYFSCNSIRVLSCGMLYERPNTYQGHQYFLRRG 66
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 32 KITDYDQENFOGKMEFTSSCPNVSESFQNVRSLSKVESGAWIGYEHTSFCGQGFILBERG 91
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY 67 DYPDYQQWGMGN----DSIRSCR-LIPQHTGTFRMRIYERDDFRGQMSHITDDCPSLQDR 121
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 92 EYPRWDWGSNGSNAYHMERLMSFRPFCSANHKESKMTIFEKENFIGRQWEISDDYPSLQAM 151
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY 122 -FHLTEVHSLNVLEGSWLYEMPSYRGROYLLR----PGEYRRLDWG--AMNAKVGSLR 174
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 152 GWFNNEVSGMKIQSGAWVCYHYLGRYGYLYLKCDHHEGDKYHREWGSHAQTSQIQSIR 211
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY 175 RV 176
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db 212 RI 213

RESULT 11
US-10-092-246-2
; Sequence 2, Application US/10092246
; Patent No. 6501314
; GENERAL INFORMATION:
; APPLICANT: The Minister of National Defence, Government of Canada
; APPLICANT: Fulton, R E
; APPLICANT: Alvi, Azhar E
; APPLICANT: Nagata, Leslie
; TITLE OF INVENTION: Cloning, Expression, Sequencing, and Functional Enhancement of Mo
; FILE REFERENCE: scfv Antibody Against Venezuelan Equine Encephalitis Virus (Vee)
; CURRENT APPLICATION NUMBER: US/10/092,246
; CURRENT FILING DATE: 2002-03-13
```

```

; APPLICANT: YAMAMOTO, Masato
; APPLICANT: YAMASHI, No. 6562599io
; APPLICANT: YAMAMOTO, Hiroko
; APPLICANT: TOHDOH, Naoki
; TITLE OF INVENTION: SINGLE-CHAIN ANTIBODY AGAINST HEPATITIS B VIRUS CORE
; TITLE OF INVENTION: PROTEIN, GENE THEROF, AND THERAPEUTIC AGENT FOR
; TITLE OF INVENTION: HEPATITIS B CONTAINING THESE
; FILE REFERENCE: 0020-4682P
; CURRENT APPLICATION NUMBER: US/09/486,814A
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Mus sp., strain: Balb/c, tissue: spleen
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(145)
; OTHER INFORMATION: Identification Method: P
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (177)..(279)
; OTHER INFORMATION: Identification Method: P
; OTHER INFORMATION:
US-09-486-814A-2

Query Match      8.7%; Score 96.5; DB 4; Length 297;
Best Local Similarity 58.3%; Pred.No.0.033;
Matches 21; Conservative 2; Mismatches 4; Indels 9; Gaps 1;

QY 172 SLRRVMDF-----YAAAGAPVPYPPDPLEPRAA 198
   |:|:|
   |||||
Db 262 SIPRIFTFCAGTKLEIKRAAGAPVPYPPDPLEPRAA 297

RESULT 14
US-09-486-814A-12
; Sequence 12, Application US/09486814A
; Patent No. 6562599
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, Masato
; APPLICANT: YAMASHI, No. 6562599io
; APPLICANT: YAMAMOTO, Hiroko
; APPLICANT: TOHDOH, Naoki
; TITLE OF INVENTION: SINGLE-CHAIN ANTIBODY AGAINST HEPATITIS B VIRUS CORE
; TITLE OF INVENTION: PROTEIN, GENE THEROF, AND THERAPEUTIC AGENT FOR
; TITLE OF INVENTION: HEPATITIS B CONTAINING THESE
; FILE REFERENCE: 0020-4682P
; CURRENT APPLICATION NUMBER: US/09/486,814A
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Targeted to Hepatitis B virus
US-09-486-814A-12

Query Match      8.7%; Score 96; DB 4; Length 18;
Best Local Similarity 100.0%; Pred.No.0.0008;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 AAAGAPVPYPPDPLEPRAA 198
   |||||
Db 1 AAAGAPVPYPPDPLEPRAA 18

RESULT 15
US-08-652-816A-43
; Sequence 43, Application US/08652816A
; Patent No. 5872215

```

GENERAL INFORMATION:
APPLICANT: Osbourn, JK
APPLICANT: Allen, DJ
APPLICANT: McCafferty, JG
TITLE OF INVENTION: Specific binding members, materials and
TITLE OF INVENTION: methods.
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,816A
FILING DATE: 23-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125579.4
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9125579.8
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 23-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525004.9
FILING DATE: 07-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9610824.6
FILING DATE: 23-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/02240
FILING DATE: 02-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/244,597
FILING DATE: 01-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/33308
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-816A-43

Query Match 8.7%; Score 96; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 181 AAAGAPVPYDPLEPRAA 198
Db 20 AAAGAPVPYDPLEPRAA 37

Search completed: July 25, 2005, 16:39:15
Job time : 27.0878 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:36:56 ; Search time 89.9696 Seconds
(without alignments)
856.072 Million cell updates/sec

Title: US-10-030-605A-19
Perfect score: 1105
Sequence: 1 AAQPMGRKFKEDRGQGH.....FYAAGAPVPVDPLEPRAA 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1741741 seqs, 388992284 residues

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	424.5	38.4	89	16	US-10-425-115-346968
2	321.5	29.1	616	15	US-10-108-260A-3095
3	317.5	28.7	781	10	US-09-866-050A-668
4	264	23.9	197	16	US-10-723-860-1946
5	253	22.9	97	9	US-09-925-298-655
6	253	22.9	97	14	US-10-102-806-655
7	211	19.1	511	15	US-10-104-047-2339
8	200	18.1	138	14	US-10-106-698-5959
9	163	14.8	83	9	US-09-864-761-44293
10	147	13.3	30	9	US-09-808-602-62
11	110	10.0	92	10	US-09-962-756-2171

12	110	10.0	92	15	US-10-253-471-2171	Sequence 2171, Ap
13	110	10.0	92	15	US-10-253-493-2171	Sequence 2171, Ap
14	109.5	9.9	194	14	US-10-112-788-10	Sequence 10, Appl
15	109.5	9.9	194	15	US-10-435-614-16	Sequence 16, Appl
16	109	9.9	60	10	US-09-962-756-2183	Sequence 2183, Ap
17	109	9.9	60	15	US-10-253-471-2183	Sequence 2183, Ap
18	109	9.9	60	15	US-10-253-493-2183	Sequence 2183, Ap
19	109	9.9	73	10	US-09-962-756-2146	Sequence 2146, Ap
20	109	9.9	73	10	US-09-962-756-2165	Sequence 2165, Ap
21	109	9.9	73	15	US-10-253-471-2146	Sequence 2146, Ap
22	109	9.9	73	15	US-10-253-471-2165	Sequence 2165, Ap
23	109	9.9	73	15	US-10-253-493-2146	Sequence 2146, Ap
24	109	9.9	73	15	US-10-253-493-2165	Sequence 2165, Ap
25	109	9.9	90	10	US-09-962-756-2142	Sequence 2142, Ap
26	109	9.9	90	10	US-09-962-756-2169	Sequence 2169, Ap
27	109	9.9	90	15	US-10-253-471-2142	Sequence 2142, Ap
28	109	9.9	90	15	US-10-253-471-2169	Sequence 2169, Ap
29	109	9.9	90	15	US-10-253-493-2142	Sequence 2142, Ap
30	109	9.9	90	15	US-10-253-493-2169	Sequence 2169, Ap
31	108.5	9.8	65	10	US-09-962-756-2170	Sequence 2170, Ap
32	108.5	9.8	65	15	US-10-253-471-2170	Sequence 2170, Ap
33	108.5	9.8	65	15	US-10-253-493-2170	Sequence 2170, Ap
34	107.5	9.7	88	10	US-09-962-756-2144	Sequence 2144, Ap
35	107.5	9.7	88	10	US-09-962-756-2168	Sequence 2168, Ap
36	107.5	9.7	88	15	US-10-253-471-2144	Sequence 2144, Ap
37	107.5	9.7	88	15	US-10-253-493-2144	Sequence 2144, Ap
38	107.5	9.7	88	15	US-10-253-471-2168	Sequence 2168, Ap
39	107.5	9.7	88	15	US-10-253-493-2168	Sequence 2168, Ap
40	107	9.7	98	10	US-09-962-756-2172	Sequence 2172, Ap
41	107	9.7	98	15	US-10-253-471-2172	Sequence 2172, Ap
42	107	9.7	98	15	US-10-253-493-2172	Sequence 2172, Ap
43	105.5	9.5	56	10	US-09-962-756-2149	Sequence 2149, Ap
44	105.5	9.5	56	10	US-09-962-756-2181	Sequence 2181, Ap
45	105.5	9.5	56	15	US-10-253-471-2149	Sequence 2149, Ap

ALIGNMENTS

RESULT 1

US-10-425-115-346968
; Sequence 346968, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 346968
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_79599C.1.pep
US-10-425-115-346968

Query Match	38.4%	Score	424.5	DB	16	Length	89
Best Local Similarity	82.2%	Pred. No.	6e-39				
Matches	74	Conservative	6	Mismatches	9	Indels	1
Qy	6	MGRKFKEDRGFGQGHYSCNSDCPNLQYFSGRCNSIRVLSGCWMLYERPNYQGHQYFLRR	65				
Db	1	MKGIITFYEDRGFGRCYECSSDCPNLQYFSGRCNSVRVDSGCWMLYERPNYQGHQYFLRR	60				
Qy	66	GDYFDYQWGMGFNDISIRSCRLIPQHTGTFR	95				
Db	61	GDYFDYQWLMGFSYIRSCRLIP-HTGSHR	89				

```
RESULT 2
US-10-108-260A-3095
; Sequence 3095, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3095
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3095

Query Match      29.1%; Score 321.5; DB 15; Length 616;
Best Local Similarity 35.2%; Pred. No. 18-25;
Matches 64; Conservative 34; Mismatches 77; Indels 7; Gaps 1;

QY 3 QPAMGRIKFKEDRGFGHYHSCNDCPNLQ-----PYFSRCNSIRVLSCGMLYEREN 55
Db 119 KGEPRAVYEPAGFGGRSWEVSRIYLNQQPDSQSPHLASVGSRLVGGCWGVEKEG 178

QY 56 YGQHQLRRGDYDPYQOMMGFNDISIRSCRLIPQHTGTFMRIRYERDDFRGQMSBITDDC 115
Db 179 FRGHQVLLGEYEPDWSHGQYDELLTSLRVIRTDGDPVAVLFRAMDPEGHGVEVSKAL 238

QY 116 PSLQDRFHTEVHSLNVLEGSWVLYEMPSYRGROYLLRPGYRRLYDGMANNAKVGSRLR 175
Db 239 PDVELVQHGFSTQAIHVLGVMVAYQEVGFSGEQYVLEKGYVRNCEDWAGNSTLASLPQ 298

QY 176 VM 177
Db 299 VL 300

RESULT 3
US-09-866-050A-668
; Sequence 668, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 668
; LENGTH: 781
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-668

Query Match      28.7%; Score 317.5; DB 10; Length 781;
Best Local Similarity 35.5%; Pred. No. 3.7e-25;
Matches 61; Conservative 33; Mismatches 71; Indels 7; Gaps 1;

QY 13 EDRGFGHYHSCNDCPNLQ-----PYFSRCNSIRVLSCGMLYERENYGHQYFLRR 65
Db 294 EAPFGQGSWEVSGDIYLNQQPDSQSPQLTSVGSRLILGGCWGVEKEGFRGHQVLLPE 353

QY 66 GDYDYOQWGMFNDSIRSCRLIPQHTGTFMRIRYERDDFRGQMSBITDDCFSLQDRPHLT 125
Db 354 GEYADWSHGQYDELLTSLRVIRTDGDPVAVLVEDMDFGHRVEVSSALPDVELAQHGP 413

QY 126 EVHSLNVLEGSWVLYEMPSYRGROYLLRPGYRRLYDGMANNAKVGSRLRRVM 177
Db 414 STQAIHVLGVMVAYERVGFSGEQVILEKGYVRNCDDWGSNGCALGSLQPVV 465

RESULT 4
US-10-723-860-1946
; Sequence 1946, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NFUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1946
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-1946

Query Match      23.9%; Score 264; DB 16; Length 197;
Best Local Similarity 34.0%; Pred. No. 4.1e-20;
Matches 65; Conservative 36; Mismatches 74; Indels 16; Gaps 7;

QY 1 AAQPAMGRIKFKEDRGFGHYHSCNDCPNL--QPYFSRCNSIRVLSCGMLYERPNYQG 58
Db 6 APGPAPASLTLDWDEDFQGRRCRLISDCANVCERGLPRVRSVKVENGVWVAFEVDFQG 65

QY 59 HQYFLRRGDYDPYQOMMGF----NDSIRSCR--LIPQHTGTFMRIRYERDDFRGQMSBIT 112
Db 66 QQFLEKGDYPRWSAWSGSSSHNSNQLLSRPVLCANHDS-RVTLFEGDNFQGGKFDLV 124

QY 113 DDCPSLQDR-PHLTEVHSLNVLEGSWVLYEMPSYRGROYLL----RPGYRRLYDGM--A 165
Db 125 DDYPSLPMGWSKADVGSLSKVSAGAWAYQYPGYRGYQVILDRDRHSGEFCTYGLGTOA 184

QY 166 MNAKVGSRLRV 176
Db 185 HTGQLQSIRRV 195

RESULT 5
US-09-925-298-655
; Sequence 655, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 655
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
```



```

:
: NAME/KEY: SITE
: LOCATION: (38)
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
:
: NAME/KEY: SITE
: LOCATION: (91)
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
: US-09-925-298-655

```

Query Match 22.9%; Score 253; DB 9; Length 97;
Best Local Similarity 51.2%; Pred. No. 2.7e-19;
Matches 41; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

[illegible]

Qy 68 YPDYQQWMGFNDISIRSLI 87
||:|:|:|:|:|:|:|:|:|:
Db 75.YPEYQRWMGLNDRLLSSXRAV 94

RESULT 6
US-10-102-806-655
; Sequence 655, Application US/10102806
; Publication No. US20030054421A1

Query Match	22.9%	Score 253;	DB 14;	Length 97;
Best Local Similarity	51.2%;	Pred. No. 2.7e-19;		
Matches 41;	Conservative 14;	Mismatches 25;	Indels 0;	Gaps 0;

Qy	8	R	I	K	F	E	D	R	G	F	Q	H	Y	S	C	N	S	D	C	P	N	L	Q	P	F	S	R	C	N	S	I	R	V	L	S	G	C	M	L	Y	E	R	P	N	Y	Q	H	Q	F	E	L	R	R	G	D	67	
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
Db	15	Q	I	T	F	E	D	K	N	F	G	R	R	V	D	C	D	C	A	D	X	H	T	Y	L	S	R	C	N	S	I	K	V	E	G	T	W	A	Y	E	R	P	N	F	A	G	Y	M	I	L	P	O	G	E	74		

Qy 68 YPDYQQWMGFNDISIRSCRLI 87
||:|:|:|:|:|:|:|:|:|:
Db 75 YPEYQRWMGLNDRLLSSXRAV 94

```

RESULT 7
US-10-104-047-2339
; Sequence 2339, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA

```

```

; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104_047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2339
; LENGTH: 511
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-104-047-2339

```

Query Match	19.1%	Score 211;	DB 15;	Length 511;
Best Local Similarity	29.0%;	Pred. No. 7.7e-14;		
Matches	47:	Conservative	32:	Mismatches 67:
				Indels 16:
				Gaps 2

QY 35 F5CNC5IRVL5GCWMLYERPNYQCHQYFLRRGDYDYOQMMGFND5IRSCLRIPOHTGT 94

DP 54 FHRIG5IRVI5GVVAYEKEHFGKCOOFL5E6GFED5NACGAL5SPIL5FRYLOANFIRS 113

[illegible]

```

RESULT 8
US-10-106-698-5959
; Sequence 5959, Application US/10106698
; Publication No. US2003010960A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer
; FILE REFERENCE: PRA00591
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5959
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5959

```

Query Match	18.1%	Score 200;	DB 14;	Length 138;
Best Local Similarity	36.0%	Pred. No. 2.3e-13;		
Matches 49:	Conservative	26: Mismatches	47: Indels	14: Gaps

Qy	54	PNYQGHGYFLRRGDYDPYQQMGF----	NDSIRSCR--LIPQHTGTFTMRIYERDDFRGQ	107
Dh	2	PNFGOGFTLRKGDYPRWSAWGSSSHNSNLI	SPRPVLCANHDS-RVTLFEGNFEQGC	60

Qy	108	MSBITDDCPSLQDR-FHLTE
Db	61	KFDLVDDYPSLPSMGWASKD
Qy	163	WG--AMNAKVGSLRRV 176
Db	121	LGTQAHGTGOLSLRRV 136

RESULT 9
US-09-864-761-44293

```

Query Match      14.8%; Score 163; DB 9; Length 83;
Best Local Similarity 38.0%; Pred. No. 1.3e-09;
Matches 30; Conservative 13; Mismatches 34; Indels 2; Gaps 1;

9 IKFKEDRFGQHYYSCNSDCNQLP--YFSRCNSIRVLSCGWMLYERYNYOCHGYFRLRG 66
1 ITLYSGKHFTGOKLRFVGDCNFDQDGRPMWVNSITHVESGAWVCNHHDFDGGOFILHGG 60

```

RESULT 13
US-10-253-493-2171
; Sequence 2171, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTUA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; CURRENT APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02

Query Match.	9.98;	Score	109.5;	DB	14;	Length	194;
Best Local Similarity	24.94;	Pred. No.	0.0023;				
Matches	59;	Conservative	24;	Mismatches	57;	Indels	97;
							Gaps 12;
Qy	1	AAAPAMGRIRKFKEDRGFQGHYYS	CNSDCPNLQVPFSCNIRVLSCGWL	YE----	----	----	----

		::::	::::	::::	::::	::::	::::
Db	16	AAQPMAMQVKQE-----	SGPEVVKPGASVKLSCKASGYIFT	SYDIDWVRQTP	63		
Qy	56	YQGHQVFLRRGDPDYQQWMGPFNDS	INRSLCLIPQHTGTFPMRIYERDD	FRQGMSEITDDC	115		
Db	64	EQG-----	LEWIGM-----	IFGEGST-----	EYNEKPKGR-ATLSVDK	96	
Qy	116	PSLQDRFHLTEVHSLNVLGSGWLYE	MPSYRGQYLLRPGCEYRYLD-WG-	-----	----	----	----
Db	97	SSSTAYMELRLTS-----	EDSAV-----	YFCARGDYRRYFLMCGQTTVT	VSSG	142	
Qy	165	-----	AMNAKVS-----	LRRVMDFYAAAGAPVPYDPLEPRAA	198		
Db	143	GGSGGGGGGGGSDIELTOSPGVKTT	QKLELR-----	AAAGAPVPYDPLEPRAA	194		

RESULT 15
US-10-435-614-16
; Sequence 16, Application US/10435614
; Publication No. US20040005647A1
; GENERAL INFORMATION:
; APPLICANT: DENARDO, SALLY J.

; APPLICANT: WINTHROP, MICHELLE D.
; APPLICANT: DENARDO, GERALD L.
; TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
; FILE REFERENCE: 309T-000220US
; CURRENT APPLICATION NUMBER: US/10/435,614
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: 10/112,788
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,721
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Mus Musculus
US-10-435-614-16

Query Match 9.9%; Score 109.5; DB 15; Length 194;
Best Local Similarity 24.9%; Pred. No. 0.0023;
Matches 59; Conservative 24; Mismatches 57; Indels 97; Gaps 12;

QY	1	AAQPMAGRIKFKEDRGFQGHYYSCNSDCPNLQYPFSRCNSIRVLSCWMLYE-----RPN	55
Db	16	AAQPMAMAQVKLQE-----SGPEWVKPGASVKLSCKASGYFTSYDIDWVRQTP	63
QY	56	YQGHQYFLRRGYPDYQQWNGFNDSIRSCRLIPQHTGTFMRRIYERDDPRGQMSITDDC	115
Db	64	EQG-----LEWIGW-----IPFGEGST-----EYNEKFKGR-ATLSVDK	96
QY	116	PSLQDRFHLTEVHSLNVLGSGWLYEMPSYRGQVLLRPGEYRRYLD-WG-----	164
Db	97	SSSTAYMELTRLTS-----EDSAV-----YFCARGDYRRYFDLWGQGTTVTVSSG	142
QY	165	-----AMNAKVGS---LRRVMDFYAAAGAPVPYDDPLEPRAA	198
Db	143	GGSGGGGGGGGSDIELTQSPGVKTGKLELKR-----AAAGAPVPYDDPLEPRAA	194

Search completed: July 25, 2005, 17:00:45
Job time : 90.3029 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:17:00 ; Search time 20.402 Seconds
(without alignments)
933.776 Million cell updates/sec

Title: US-10-030-605A-19
Perfect score: 1105
Sequence: 1 AAQPMGRKPKEDRGFOGH.....FYAAGAPVPYDPLEPRAA 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	943	85.3	175	1	gamma-crystallin I
2	881	79.7	175	2	gamma-crystallin 1
3	864	78.2	175	1	gamma-B-crystallin
4	830	75.1	175	1	gamma-crystallin 1
5	783.5	70.9	174	2	gamma-crystallin 2
6	778.5	70.5	174	1	gamma-crystallin 2
7	778.5	70.5	174	2	gamma-crystallin 2
8	776.5	70.3	174	2	gamma-C-crystallin
9	775.5	70.2	174	2	gamma-crystallin 2
10	774.5	70.1	174	2	gamma-crystallin 1
11	773.5	70.0	173	1	gamma-crystallin 1
12	770.5	69.7	173	2	gamma-crystallin 1
13	770.5	69.7	174	1	gamma-crystallin 1
14	767.5	69.5	174	1	gamma-crystallin 1
15	765.5	69.3	174	1	gamma-crystallin 4
16	764.5	69.2	174	2	gamma-crystallin 4
17	762.5	69.0	174	2	gamma-crystallin 4
18	756.5	68.5	174	2	gamma-crystallin 4
19	752.5	68.1	174	2	gamma-crystallin -
20	740.5	67.0	170	1	gamma-crystallin 2
21	726.5	65.7	174	1	gamma-crystallin 2
22	723.5	65.5	174	2	gamma-D-crystallin
23	715.5	64.8	174	2	gamma-crystallin 5
24	713.5	64.6	157	2	gamma-crystallin I
25	686	62.1	176	2	gamma-crystallin M
26	684.5	61.9	159	2	gamma-F-crystallin
27	682.5	61.8	151	2	gamma-C-crystallin
28	656	59.4	173	2	gamma-crystallin S
29	645	58.4	175	2	gamma3-crystallin

30	639	57.8	168	1	gamma-crystallin I
31	639	57.8	169	2	gamma-2-crystallin
32	631	57.1	175	2	gamma2-crystallin
33	627	56.7	177	2	gamma-M1-1 crystal
34	621	56.2	175	2	gamma1-crystallin
35	616	55.7	177	2	gamma-M1-2 crystal
36	613	55.5	172	2	gamma4-crystallin
37	610	55.2	173	2	gamma-crystallin S
38	561	50.8	133	1	gamma-crystallin I
39	554.5	50.2	174	2	gamma-s-crystallin
40	545.5	49.4	174	2	gamma-crystallin m
41	543.5	49.2	177	2	gamma-crystallin m
42	538	48.7	173	1	gamma-crystallin m
43	537.5	48.6	178	1	gamma-s-crystallin
44	532.5	48.2	178	2	gamma-crystallin M
45	531.5	48.1	175	2	gamma-crystallin M

ALIGNMENTS

RESULT 1

CYBOG

gamma-crystallin II - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 28-Feb-1986 #sequence revision 03-Feb-1994 #text change 09-Jul-2004

C:Accession: A29655; A02928; A90263; A60890; A60815; A02929; S04265

R:Hay, R.E.; Woods, W.D.; Church, R.L.; Petrasch, J.M.

Biochem. Biophys. Res. Commun. 146, 332-338, 1987

A:Title: cDNA clones encoding bovine gamma-crystallins.

A:Reference number: A90133; MUID:8720760; PMID:3606621

A:Accession: A29655

A:Molecule type: mRNA

A:Residues: 1-175 <HAY>

A:Cross-references: UNIPROT:P02526; GB:M16894; NID:gl62918; PID:AAA30476.1; PID:gl62919

R:Bhat, S.P.; Spector, A.

DNA 3, 287-295, 1984

A:Title: Complete nucleotide sequence of a cDNA derived from calf lens gamma-crystallin

A:Reference number: A02928; MUID:85026663; PMID:6092016

A:Accession: A02928

A:Molecule type: mRNA

A:Residues: 2-119,'S',121-175 <BHA>

A:Cross-references: GB:K02112; GB:X01036; NID:gl62916; PID:AAA30475.1; PID:gl62917

A:Note: Initiator Met not shown

R:Croft, L.R.

Biochem. J. 128, 961-970, 1972

A:Title: The amino acid sequence of gamma-crystallin (fraction II) from calf lens.

A:Reference number: A90263; MUID:73054483; PMID:4674126

A:Accession: A90263

A:Molecule type: protein

A:Residues: 2-17,'Q',19-21,'NN',23,'LOP',28-39,'VHSL',45-46,'MLQ',48-49,'D',51,53-54,56,

A:Note: Portions of this sequence were assigned by composition rather than by direct sequencing.

R:Chou, S.H.; Azari, P.; Hamel, M.E.

J. Protein Chem. 7, 67-80, 1988

A:Title: Physicochemical characterization of gamma-crystallins from bovine lens-hydroxy

A:Reference number: A60890; MUID:89351571; PMID:3255364

A:Accession: A60890

A:Molecule type: protein

A:Residues: 2-26 <CHI>

R:McDermott, M.J.; Gawinowicz-Kolke, M.A.; Chiesa, R.; Spector, A.

Arch. Biochem. Biophys. 262, 609-619, 1988

A:Title: The disulfide content of calf gamma-crystallin.

A:Reference number: A60815; MUID:88208422; PMID:3364984

A:Accession: A60815

A:Molecule type: protein

A:Residues: 2-26 <MCD>

R:Blundell, T.; Lindley, P.; Miller, L.; Moss, D.; Slingsby, C.; Tickle, I.; Turnell, B.

Nature 289, 771-777, 1981

A:Title: The molecular structure and stability of the eye lens: x-ray analysis of gamma-

A:Reference number: A93247; MUID:8112311; PMID:7464942

A:Contents: annotation; X-ray crystallography, 2.6 angstroms

R:White, H.E.; Driessen, H.P.C.; Slingsby, C.; Moss, D.S.; Lindley, P.F.

J. Mol. Biol. 207, 217-235, 1989

A>Title: Extensive intragenic sequence homology in two distinct rat lens gamma-crystallin
A:Reference number: A93934; MUID:83091061; PMID:6294661
A:Accession: A02930
A:Molecule type: mRNA
A:Residues: 1-173 <MOO>
A:Cross-references: UNIPROT:P02528; GB:J00716; NID:G203634; PIDN:AAA40987.1; PID:G203635
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens
F:1-39/Domain: crystallin repeat <GK1>
F:40-82/Domain: crystallin repeat <GK2>
F:87-127/Domain: crystallin repeat <GK3>
F:128-167/Domain: crystallin repeat <GK4>

Query Match 70.0%; Score 773.5; DB 1; Length 173;
Best Local Similarity 77.0%; Pred. No. 6.6e-66;
Matches 134; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGCMVLYERPNYQGHQYFLRRG 66
DB 1 GKITTFYEDRGFGQHRHYECSTDSHNLQPYFSRCNSVRVDSGCMVLYEQPNFTGCQYFLRRG 60
QY 67 DYPDYQQWGMFNDIRSCLIPQHTGTFMRRIYERDDFRGOMSEITDDCPSLQDRPHLITE 126
DB 61 DYPDYQQWGMFSDSVRSCLIP-HSSSHRIYERDYRGQWVEITDDCPHLQDRPHFSD 119
QY 127 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 180
DB 120 FHSFHVMEGYWVLYEMPENYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRIMDFY 173

RESULT 12
S04266
gamma-crystallin IV - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
C:Accession: S04266; C60890
R:White, H.E.; Driessen, H.P.C.; Slingsby, C.; Moss, D.S.; Lindley, P.F.
J. Mol. Biol. 207, 217-235, 1989
A>Title: Packing interactions in the eye-lens. Structural analysis, internal symmetry and
A:Reference number: S04265; MUID:89293855; PMID:2738925
A:Accession: S04266
A:Molecule type: protein
A:Residues: 1-173 <WHI>
R:Chiou, S.H.; Azari, P.; Himmel, M.E.
J. Protein Chem. 7, 67-80, 1988
A>Title: Physicochemical characterization of gamma-crystallins from bovine lens-hydrodynam
A:Reference number: A60890; MUID:89351571; PMID:3255364
A:Accession: C60890
A:Molecule type: protein
A:Note: 14-His, 15-Arg, 22-Cys, 23-Pro, and 24-Asp were also found and may indicate cont
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens

Query Match 69.7%; Score 770.5; DB 2; Length 173;
Best Local Similarity 77.0%; Pred. No. 1.3e-65;
Matches 134; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGCMVLYERPNYQGHQYFLRRG 66
DB 1 GKITTFYEDRGFGQHRHYECSSDSHNLQPYFSRCNSIRVDSGCMVLYEQPNFTGCQYFLRRG 60
QY 67 DYPDYQQWGMFNDIRSCLIPQHTGTFMRRIYERDDFRGOMSEITDDCPSLQDRPHLITE 126
DB 61 DYPDYQQWGMFSDSVRSCLIP-HTSSSHRLIYERDYRGQWVEITDDCSSLQDRPHFSD 119
QY 127 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 180
DB 120 IHSFHVMEGYWVLYEMPENYRGQYLLRPGDYRRYLDWGAMNAKVGSLRRVMDFY 173

RESULT 13
CYMSG1

gamma-crystallin 1 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
A:Accession: A02932
R:Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Tretton, J.A.; Gold, R.J.M.; Tsui
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A>Title: gamma-crystallin family of the mouse lens: structural and evolutionary relations
A:Reference number: A94021; MUID:85088487; PMID:6096855
A:Accession: A02932
A:Molecule type: mRNA
A:Residues: 1-174 <BRE>
A:Cross-references: UNIPROT:P04342; GB:K02583; NID:G192778; PIDN:AAA37475.1; PID:G309197
C:Comment: There are at least seven different gamma crystallins identified in mouse lens
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens
F:2-40/Domain: crystallin repeat <GK1>
F:41-83/Domain: crystallin repeat <GK2>
F:88-128/Domain: crystallin repeat <GK3>
F:129-168/Domain: crystallin repeat <GK4>

Query Match 69.7%; Score 770.5; DB 1; Length 174;
Best Local Similarity 76.6%; Pred. No. 1.3e-65;
Matches 134; Conservative 21; Mismatches 19; Indels 1; Gaps 1;

QY 6 MGRIKFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGCMVLYERPNYQGHQYFLRR 65
DB 1 MGKITTFYEDRGFGQHRHYECSTDSHNLQPYFSRCNSVRVDSGCMVLYEQPNFAGCQYFLRR 60
QY 66 GDYPDYQQWGMFNDIRSCLIPQHTGTFMRRIYERDDFRGOMSEITDDCPSLQDRPHLT 125
DB 61 GDYPDYQQWGMFSDSVRSCLIP-HAGSHRLIYERDYRGQWVEITDDCPSLQDRPHFN 119
QY 126 EVHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 180
DB 120 EYLSNLVLEGSWVLYDMTYNRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 174

RESULT 14
JS0596
gamma-E-crystallin - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Nov-1997
A:Accession: JS0596; S18835
R:Graw, J.; Coban, L.; Liebststein, A.; Werner, T.
Gene 104, 265-270, 1991
A>Title: Murine gamma E-crystallin is distinct from murine gamma 2-crystallin.
A:Reference number: JS0596; MUID:92009223; PMID:1916296
A:Accession: JS0596
A:Molecule type: DNA
A:Residues: 1-174 <GRA>
A:Cross-references: EMBL:X57855
A:Note: this sequence differs two residues from that of rat
C:Genetics:
A:Gene: gamma-E-cry
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication

Query Match 69.5%; Score 767.5; DB 2; Length 174;
Best Local Similarity 76.0%; Pred. No. 2.5e-65;
Matches 133; Conservative 22; Mismatches 19; Indels 1; Gaps 1;

QY 6 MGRIKFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGCMVLYERPNYQGHQYFLRR 65
DB 1 MGKITTFYEDRGFGQHRHYECSTDSHNLQPYFSRCNSVRVDSGCMVLYEQPNFTGCQYFLRR 60
QY 66 GDYPDYQQWGMFNDIRSCLIPQHTGTFMRRIYERDDFRGOMSEITDDCPSLQDRPHLT 125
DB 61 GDYPDYQQWGMFSDSVRSCLIP-HSSSHRIYERDYRGQWVEITDDCSHLQDRPHFS 119
QY 126 EVHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 180
DB 120 DFHSFHVMEGYWVLYEMPENYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRIMDFY 174

RESULT 15

CYMSG4
gamma-crystallin 4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A02935; I49613
R:Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Tretton, J.A.; Gold, R.J.M.; Tsui
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A:Title: gamma-crystallin family of the mouse lens: structural and evolutionary relations
A:Reference number: A94021; MUID:85088487; PMID:6096855
A:Accession: A02935
A:Molecule type: mRNA
A:Residues: 1-174 <BRE>
A:Cross-references: UNIPROT:P04345
A:Note: the authors translated the codon ATC for residue 36 as Val
R:Lok, S.; Tsui, L.C.; Shinohara, T.; Piatigorsky, J.; Gold, R.; Breitman, M.
Nucleic Acids Res. 12, 4517-4529, 1984
A:Title: Analysis of the mouse gamma-crystallin gene family: assignment of multiple cDNA
A:Reference number: I48353; MUID:84247318; PMID:8336674
A:Accession: I49613
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-174 <RES>
A:Cross-references: GB:K02587; NID:912771; PIDN:AAA37473.1; PID:g387135
C:Comment: There are at least seven different gamma crystallins identified in mouse lens
C:Genetics:
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens
F:2-40/Domain: crystallin repeat <GK1>
F:41-83/Domain: crystallin repeat <GK2>
F:88-128/Domain: crystallin repeat <GK3>
F:129-168/Domain: crystallin repeat <GK4>

Query Match 69.3%; Score 765.5; DB 1; Length 174;
Best Local Similarity 77.1%; Pred. No. 3.8e-65;
Matches 135; Conservative 19; Mismatches 20; Indels 1; Gaps 1;
QY 6 MGRIRKFEDRGFGQHYYSCNSDCPNLQPYFSRCNSIRVLSCGCMILYERPNTYQGHQYFLRR 65
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MGKITFYEDRGFGQRCYECSSDCPNLQTYFSRCNSIRVDSGCMILYERPNTYQGYFLRR 60
QY 66 GDYDYOQWMGNDNSIRSCRLIPQHTGTFRMIRYERDDFRGQMSBITDDCPSLQDRFHLT 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 GDYDYOQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELMDDCSCIHDRFRLH 119
QY 126 EVHSLNVLEGSWLYEMPSPYRGROYLLRPGEYRRLDWMGNNAKVGSLRRVMDFY 180
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 120 EYSHVLEGCWLYEMPNTYRGROYLLRPGDTRRYTHDWMADKVGSLRRVMDLY 174

Search completed: July 25, 2005, 16:37:51
Job time : 21.402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:15:10 ; Search time 91.6419 Seconds
(without alignments)
1106.390 Million cell updates/sec

Title: US-10-030-605A-19

Perfect score: 1105

Sequence: 1 AAQPMGRKPKEDRGFGH.....FYAAGAPVPVDPLEPRAA 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	84.9	174	1	CRGB_BOVIN
2	876	79.3	174	1	CRGB_RAT
3	859	77.7	174	1	CRGB_MOUSE
4	856	77.5	175	2	O6PH7
5	825	74.7	174	1	CRGB_HUMAN
6	801	72.5	174	1	CRGA_BOVIN
7	778.5	70.5	173	1	CRGD_BOVIN
8	778.5	70.5	173	1	CRGD_RAT
9	776.5	70.3	174	2	O6PGIO
10	773.5	70.0	173	1	CRGE_RAT
11	772.5	69.9	173	1	CRGF_BOVIN
12	771.5	69.8	173	1	CRGC_RAT
13	769.5	69.6	173	1	CRGA_RAT
14	765.5	69.3	173	1	CRGD_MOUSE
15	762.5	69.0	173	1	CRGB_MOUSE
16	760.5	68.8	173	1	CRGA_MOUSE
17	757.5	68.6	173	1	CRGF_RAT
18	752.5	68.1	173	1	CRGF_MOUSE
19	751.5	68.0	173	1	CRGE_BOVIN
20	747.5	67.6	173	1	CRGD_MOUSE
21	722.5	65.4	173	1	CRGD_HUMAN
22	721.5	65.3	173	1	CRGC_HUMAN
23	710.5	64.3	173	1	CRGA_HUMAN
24	689	62.4	174	2	O66KM8
25	681	61.6	175	1	CRG2_CHICO
26	671	60.7	175	2	O6DKC9
27	654	59.2	175	2	O7SZA6
28	651	58.9	172	1	CRBS_CHICO
29	649	58.7	175	2	O6DJC9
30	647	58.6	175	2	O6DER7
31	645	58.4	175	1	CRG3_XENLA

32	644	58.3	174	2	O9PSV5	O9PSV5 xenopus lae
33	643	58.2	174	2	O68ES5	O68ES5 xenopus lae
34	642	58.1	175	2	O66KW2	O66KW2 xenopus lae
35	641	58.0	175	2	O8QFU3	O8QFU3 cynops pyrr
36	641	58.0	189	2	O66KR5	O66KR5 xenopus lae
37	639	57.8	169	1	CRG2_RANTE	P02531 rana tempor
38	635	57.5	175	2	O66LI5	O66LI5 xenopus lae
39	635	57.5	177	2	O36LI5	O36LI5 xenopus lae
40	632	57.2	175	2	O66KU7	O66KU7 xenopus lae
41	631	57.1	175	1	CRG2_XENLA	O91724 xenopus lae
42	627	56.7	177	1	CRG1_XENLA	O91320 rana catesb
43	621	56.2	175	1	CRG1_XENLA	O62554 xenopus lae
44	616	55.7	177	1	CRG2_RANCA	O91321 rana catesb
45	610	55.2	172	1	CRG4_XENLA	P55941 xenopus lae

ALIGNMENTS

RESULT 1

ID	CRGB_BOVIN	STANDARD	PRT	174 AA
AC	P02526;			
DT	21-JUN-1986 (Rel. 01, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	Gamma crystallin B (Gamma crystallin II).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
CC	Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87270760; PubMed=3606621;			
RA	Hay R.E., Woods W.D., Church R.L., Petrasch J.M.;			
RT	"cDNA clones encoding bovine gamma-crystallins.";			
RL	Biochem. Biophys. Res. Commun. 146:332-338(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lens;			
RX	MEDLINE=85026663; PubMed=6092016;			
RA	Bhat S.P., Spector A.;			
RT	"Complete nucleotide sequence of a cDNA derived from calf lens gamma-crystallin mRNA: presence of Alu I-like DNA sequences.";			
RL	DNA 3:287-295(1984).			
RN	[3]			
RP	SEQUENCE OF 1-25.			
RX	PubMed=3255364;			
RA	Chieu S.H., Azari P., Himmel M.E.;			
RT	"Physicochemical characterization of gamma-crystallins from bovine lens -- hydrodynamic and biochemical properties.";			
RL	J. Protein Chem. 7:67-80(1988).			
RN	[4]			
RP	SEQUENCE OF 1-25, AND DISULFIDE BOND.			
RX	PubMed=3364984;			
RA	McDermott M.J., Gawinowicz-Kolks M.A., Chiesa R., Spector A.;			
RT	"The disulfide content of calf gamma-crystallin.";			
RL	Arch. Biochem. Biophys. 262:609-619(1988).			
RN	[5]			
RP	X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS), AND REVISIONS.			
RX	MEDLINE=81123111; PubMed=7464942;			
RA	Blundell T.L., Lindley P., Miller L., Moss D., Slingsby C., Tickle I.,			
RT	Turnell B., Wistow G.;			
RL	"The molecular structure and stability of the eye lens: X-ray analysis of gamma-crystallin II.";			
RN	Nature 289:771-777(1981).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RX	MEDLINE=84036195; PubMed=6631960;			
RA	Wistow G., Turnell B., Summers L., Slingsby C., Moss D., Miller L.,			
RL	Lindley P., Blundell T.L.;			
RT	"X-ray analysis of the eye lens protein gamma-II crystallin at 1.9-A			

```

RT resolution.";
RL J. Mol. Biol. 170:175-202(1983).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.47 ANGSTROMS).
RX PubMed=15299528;
RA Najmudin S., Nalini V., Dreissen H.P.C., Slingsby C., Blundell T.L.,
RA Moss D.S., Lindley P.F.;
RT "Structure of the bovine eye lens protein gammaB (gammaII)-crystallin
RT at 1.47 A.";
RL Acta Crystallogr. D 49:223-233(1993).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).
RX PubMed=15299624;
RA Kumaraswamy V.S., Lindley P.F., Slingsby C., Glover I.D.;
RT "An eye lens protein-water structure: 1.2-A resolution structure of
RT gammaB-crystallin at 150 K.";
RL Acta Crystallogr. D 52:611-622(1996).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX MEDLINE=98200466; PubMed=9541393;
RA Palme S., Jaenicke R., Slingsby C.;
RT "X-ray structures of three interface mutants of gammaB-crystallin from
RT bovine eye lens.";
RL Protein Sci. 7:611-618(1998).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT ALA-56.
RX MEDLINE=98308013; PubMed=9642083; DOI=10.1006/jmbi.1998.1850;
RA Palme S., Jaenicke R., Slingsby C.;
RT "Unusual domain pairing in a mutant of bovine lens gammaB-
RT crystallin.";
RL J. Mol. Biol. 279:1053-1059(1998).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin 'Greek key' domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01036; CAA25518.1; -;
DR EMBL; M16894; AAA30476.1; -;
DR PIR; A29655; CYBOG.
DR PDB; 1AMM; X-ray; @=1-174.
DR PDB; 1DSL; X-ray; @=87-174.
DR PDB; 1GAM; X-ray; A/B=87-172.
DR PDB; 1GCS; X-ray; @=1-174.
DR PDB; 1I51; X-ray; A=1-174.
DR PDB; 4GCR; X-ray; @=1-174.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; Crystallin_SF.
DR Pfam; PF00303; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALbg; 2.
DR PROSITE; PS05915; CRYSTALLIN BETAGAMMA; 4.
KW 3D-structure; Direct protein sequencing; Eye lens protein;
KW Multigene family; Repeat.
FT INIT MET 0 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 87 Connecting peptide.
FT DOMAIN 88 128 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 129 171 Beta/gamma crystallin 'Greek key' 4.
FT DISULFID 18 22
FT CONFLICT 119 119 T -> S (in Ref. 2).
FT STRAND 2 8
FT HELIX 9 11

```

```

FT STRAND 12 18
FT STRAND 22 22
FT TURN 26 28
FT STRAND 34 39
FT STRAND 42 47
FT TURN 48 50
FT STRAND 51 57
FT STRAND 60 62
FT HELIX 65 68
FT TURN 69 69
FT STRAND 77 80
FT STRAND 89 95
FT TURN 96 98
FT STRAND 99 105
FT STRAND 109 109
FT HELIX 112 116
FT STRAND 120 120
FT STRAND 123 128
FT STRAND 131 136
FT TURN 137 139
FT STRAND 140 146
FT STRAND 149 151
FT HELIX 154 157
FT TURN 158 158
FT STRAND 163 163
FT STRAND 166 169
SQ SEQUENCE 174 AA; 20965 MW; 8E404878CA2150A3 CRC64;

Query Match      84.9%; Score 938; DB 1; Length 174;
Best Local Similarity 96.0%; Pred. No. 1.6e-80;
Matches 167; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 GRIFKEDRGFGQHYHSCNDCPNLQPYFSCNSIRVLSCGMWLYERPNYQGHQYFLRRG 66
DB 1 GKITYEDRGFGQHYHSCNDCPNLQPYFSCNSIRVLSCGMWLYERPNYQGHQYFLRRG 60

QY 67 DYPDYOQWGMFNDSIRSCRLIPQHTGTFRMIRYERDDPRGOMSETDDCPSLQDRFHLE 126
DB 61 DYPDYOQWGMFNDSIRSCRLIPQHTGTFRMIRYERDDPRGOMSETDDCPSLQDRFHLE 120

QY 127 VHSNLVLEGSWVLYEMPFSYGRQYLLRPFGEYRRYLDWGAMNAKVGSLLRRVMDFY 180
DB 121 VHSNLVLEGSWVLYEMPFSYGRQYLLRPFGEYRRYLDWGAMNAKVGSLLRRVMDFY 174

RESULT 2
CRGB RAT STANDARD; PRT; 174 AA.
AC P10066;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin B (Gamma crystallin 1-2).
GN Name=Crygb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90236314; PubMed=1970548; DOI=10.1016/0378-1119(90)90306-C;
RA den Dunnen J.T., van Neck J.W., Creemers F.P.M., Lubsen N.H.,
RA Schoenmakers J.G.G.;
RT "Isolation and characterization of the rat glutamine synthetase-
RT encoding gene.";
RL Gene 87:225-232(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060933; PubMed=3783678;
RA den Dunnen J.T., Moormann R.J.M., Lubsen N.H., Schoenmakers J.G.G.;
RT "Concerted and divergent evolution within the rat gamma-crystallin
RT gene family.";
RL J. Mol. Biol. 189:37-46(1986).

```



```

Db      1 GKITYEDRAFGSRVECTTDCPNLQPYFSRCNSIRVSGCWMYIYERNYQCHQYFLRRG 60
QY      67 DYPDYQWGMFNDIRSCLIPQHTGTFMRRIYERDDFRGQWSEITDDCPSLQDRFHLTE 126
Db      61 EYPDYQWGMGLSDIRSCLIPPHSGAVRMKIYDRDELRGQWSELTDCLSLVQDRFHLTE 120
QY      127 VHSNLVLEGSWLYEMPSYRGQYLLRPGCEYRYLLDGMNNAKVGSLRRVMDFY 180
Db      121 IHSNLVLEGSWLYEMPNYRGQYLLRPGCEYRYLLDGMNNAKVGSLRRVMDLY 174

RESULT 6
CRGA_BOVIN
ID CRGA_BOVIN STANDARD; PRT; 174 AA.
AC P02527;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Gamma crystallin A (Gamma crystallin IVB).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=73054483; PubMed=4674126;
RA Croft L.R.;
RT "The amino acid sequence of gamma-crystallin (fraction II) from calf
RT lens."
RL Biochem. J. 128:961-970(1972).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
DR HSP; P02526; IDSL.
DR InterPro; IPR01024; G.crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR SMART; SM00247; XTALbg; 2.
DR PROSITE; PS50915; CRYSTALLIN BETAGAMMA; 4.
DR Direct protein sequencing; Eye lens protein; Multigene family; Repeat.
KW DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 87 Connecting peptide.
FT DOMAIN 88 128 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 129 171 Beta/gamma crystallin 'Greek key' 4.
SQ SEQUENCE 174 AA; 21003 MW; E35093A105E2CEEFCRC64;

Query Match 72.5%; Score 801; DB 1; Length 174;
Best Local Similarity 82.8%; Pred. No. 1.4e-67;
Matches 144; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

QY      7 GRIKEDRGFGCHYVSCNDCPNLQPYFSRCNSIRVSGCWMYIYERNYQCHQYFLRRG 66
Db      1 GKITYEDRGFGCHYVSCNDCPNLQPYFSRCNSIRVSGCWMYIYERNYQCHQYFLRRG 60
QY      67 DYPDYQWGMFNDIRSCLIPQHTGTFMRRIYERDDFRGQWSEITDDCPSLQDRFHLTE 126
Db      61 NYPQYQWGMFNDIRSCLIPQHTGTFMRRIYERDDFRGQWSEITDDCPSLQDRFHLTE 120
QY      127 VHSNLVLEGSWLYEMPSYRGQYLLRPGCEYRYLLDGMNNAKVGSLRRVMDFY 180
Db      121 VNSRVLEGSWLYEMPSYRGQYLLRPGCEYRYLLDGMNNAKVGSLRRVMDFY 174

RESULT 7
CRGD_BOVIN
ID CRGD_BOVIN STANDARD; PRT; 173 AA.
AC P08209; Q28089;

```

```

DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin D (Gamma crystallin IIIB).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE=Lens;
RX MEDLINE=95010404; PubMed=7925695; DOI=10.1006/exer.1994.1052;
RA Hay R.E., Andley U.P., Petrash J.M.;
RT "Expression of recombinant bovine gamma B-, gamma C- and gamma D-
RT crystallins and correlation with native proteins."
RL Exp. Eye Res. 58:573-584(1994).
RN [2]
RP SEQUENCE OF 1-156 FROM N.A.
RX TISSUE=Lens;
RX MEDLINE=87270760; PubMed=3606621;
RA Hay R.E., Woods W.D., Church R.L., Petrash J.M.;
RT "cDNA clones encoding bovine gamma-crystallins."
RL Biochem. Biophys. Res. Commun. 146:332-338(1987).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RX PubMed=15299634;
RA Chirgadze Y.N., Driessen H.P.C., Wright G., Slingsby C., Hay R.E.,
RA Lindley P.F.;
RT "Structure of bovine eye lens gammaD (gammaIIb)-crystallin at 1.95
RT A."
RL Acta Crystallogr. D 52:712-721(1996).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L27070; AAA98995.1; -.
EMBL; M16895; AAB59282.1; -.
PDB; IELP; X-ray; A/B=1-173.
InterPro; IPR001064; Crystallin.
InterPro; IPR011024; G.crystallin_SF.
Pfam; PF00030; Crystall; 2.
PRINTS; SM00247; XTALbg; 2.
SMART; SM00247; XTALbg; 2.
PROSITE; PS50915; CRYSTALLIN BETAGAMMA; 4.
3D-structure; Direct protein sequencing; Eye lens protein;
Multigene family; Repeat.
INIT_MET 0 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT CONFLICT 75 75 V -> I (in Ref. 2).
FT STRAND 2 8
FT HELIX 9 11
FT STRAND 12 18
FT STRAND 22 22
FT TURN 26 28
FT STRAND 34 39
FT STRAND 42 45
FT TURN 49 50

```

```

FT STRAND 54 57
FT STRAND 60 62
FT HELIX 65 68
FT TURN 69 69
FT STRAND 77 80
FT STRAND 88 94
FT TURN 95 97
FT STRAND 98 104
FT STRAND 108 108
FT HELIX 111 114
FT TURN 115 115
FT STRAND 119 119
FT STRAND 122 127
FT STRAND 130 135
FT TURN 136 138
FT STRAND 139 145
FT STRAND 148 150
FT HELIX 153 156
FT TURN 157 157
FT STRAND 162 162
FT STRAND 165 168
SQ SEQUENCE 173 AA; 20735 MW; D9D853EBB3B3F7B5F CRC64;

Query Match 70.5%; Score 778.5; DB 1; Length 173;
Best Local Similarity 78.2%; Pred. No. 1.8e-65;
Matches 136; Conservative 17; Mismatches 20; Indels 1; Gaps 1;

QY 7 GRIKFEDRGFGQHYHSCNDCPNLQPYFSCNSIRVLSCGWMLYERNYQHQYFLRRG 66
Db 1 GKITFYEDRGFGQHRHYECSDHSLNQLQPYFSCNSVRVDSGCMWLYEQPNYLGQYFLRRG 60

QY 67 DYPDYQQWGMGENDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRFHLTE 126
Db 61 DYPDYQQWGLNDSVRSCLIP-HAGSHRLRLYEREDYRGQMIEITEDCSSLQDRFHFNE 119

QY 127 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 180
Db 120 IHSNLVLEGSWVLYELPNYRGQYLLRPGEYRRYHDWGAMNAKVGSLRRVIDY 173

RESULT 8
CRGD RAT
ID _CRGD RAT STANDARD; PRT; 173 AA.
AC P10067;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin D (Gamma crystallin 2-2).
GN Name=Crygd;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90236314; PubMed=1970548; DOI=10.1016/0378-1119(90)90306-C;
RA den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA Schoenmakers J.G.G.;
RT "Isolation and characterization of the rat glutamine synthetase-
RT encoding gene.";
RL Gene 87:225-232(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060933; PubMed=3783678;
RA den Dunnen J.T., Moormann R.J.M., Lubsen N.H., Schoenmakers J.G.G.;
RT "Concerted and divergent evolution within the rat gamma-crystallin
RT gene family.";
RL J. Mol. Biol. 189:37-46(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Lens;
RX MEDLINE=95151850; PubMed=7849105; DOI=10.1016/0300-9084(94)90115-5;
RA Ooki K., Amuro N., Shimizu Y., Okazaki T.;
```

```

RT "High level expression of rat gamma-D-crystallin in Escherichia
RT coli.";
RL Biochimie 76:398-403(1994).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- MISCELLANEOUS: There are six different gamma crystallins
CC identified in rat lens.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M19359; AAA40984.1; -.
DR EMBL; X57169; CAA40458.1; -.
DR PIR; D24060; D24060.
DR HSSP; P02528; 1A5D.
DR RGD; 2422; Crygd.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G-Crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALbg; 2.
DR PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW Eye lens protein; Multigene family; Repeat.
FT INIT_MET 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
SQ SEQUENCE 173 AA; 20972 MW; 930B2D42D80F4117 CRC64;

Query Match 70.5%; Score 778.5; DB 1; Length 173;
Best Local Similarity 78.2%; Pred. No. 1.8e-65;
Matches 136; Conservative 19; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFEDRGFGQHYHSCNDCPNLQPYFSCNSIRVLSCGWMLYERNYQHQYFLRRG 66
Db 1 GKITFYEDRGFGQHRHYECSDHSLNQLQPYFSCNSVRVDSGCMWLYEQPNYLGQYFLRRG 60

QY 67 DYPDYQQWGMGENDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRFHLTE 126
Db 61 DYPDYQQWGMGFSVRSCLIP-HAGSHRIKLYEREDYRGQNVFTEDCPSLQDRFHFNE 119

QY 127 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 180
Db 120 IYSLNLVLEGSWVLYEMTYNRGRQYLLRPGEYRRYHDWGAMNARVGSLLRRVMDFY 173

RESULT 9
Q6PGIO
ID Q6PGIO PRELIMINARY; PRT; 174 AA.
AC Q6PGIO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Crystallin, gamma D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
```


RA Strausberg R.L., Feinsold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Crystallins are the dominant structural components of
the vertebrate eye lens (By similarity).
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
similar Greek key motifs (By similarity).
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
DR EMBL; BC057013; AAH57013.1; -.
DR HSSP; P02528; 1A5D.
DR InterPro; IPR001064; G crystallin.
DR InterPro; IPR011024; G crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALbg; 2.
DR PROSITE; PS0915; CRYSTALLIN BETAGAMMA; 4.
SQ SEQUENCE 174 AA; 21118 MW; 1AFC29C0269B35A1 CRC64;

Query Match 70.3%; Score 776.5; DB 2; Length 174;
Best Local Similarity 77.1%; Pred. No. 2.8e-65;
Matches 135; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 6 MGRIKPKEDRGFGHYHSCNDCPNLQPYFSCNCSIRVLSCGMWLYERNYQGHQYFLRR 65
DB 1 MGKITFYDRGFGGRHYEGCTHSNLSQPYFSCNCSVRVDSGCMWLYEQNFQCGQFLRR 60

QY 66 GYDPYQQMGFNDSIRSLIPQHTGTFMRIRYERDDFRGQWSEITDCPSLQDRFHLT 125
DB 61 GYDPYQQMGFSDSVRSCLIP-HAGSHRILRYEREYRGQMIEFTDCPSLQDRFHN 119

QY 126 EYHSLNVLGSGWLYEMPVSRGQYLLRGEYRYLDWGMNAKVGSLRRVMDFY 180
DB 120 EYSLNVLGSGWLYDMTYNRGQYLLRGEYRRYHDWGMNARVGLRRVMDFY 174

RESULT 10
CRGE_RAT STANDARD; PRT; 173 AA.
AC P02528;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin B (Gamma crystallin 3-1) (Gamma-2).
GN Names=Cryge;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90236314; PubMed=1970548; DOI=10.1016/0378-1119(90)90306-C;

RA den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA Schoenmakers J.G.G.;
RT "Isolation and characterization of the rat glutamine synthetase-
encoding gene.";
RL Gene 87:225-232(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83091061; PubMed=6294661;
RA Moormann R.J.M., den Dunnen J.T., Bloemendal H., Schoenmakers J.G.G.;
RT "Extensive intragenic sequence homology in two distinct rat lens
gamma-crystallin cDNAs suggests duplications of a primordial gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6876-6880(1982).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=84114867; PubMed=6319707;
RA Moormann R.J.M., den Dunnen J.T., Mulleners L., Andreoli P.,
RA Bloemendal H., Schoenmakers J.G.G.;
RT "Strict co-linearity of genetic and protein folding domains in an
intragenically duplicated rat lens gamma-crystallin gene.";
RL J. Mol. Biol. 171:353-368(1983).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RC TISSUE=Lens;
RX MEDLINE=98035705; PubMed=9367641; DOI=10.1006/exer.1997.0368;
RA Norledge B.V., Hay R.E., Bateman O.A., Slingsby C., Driessen H.P.C.;
RT "Towards a molecular understanding of phase separation in the lens: a
comparison of the X-ray structures of two high Tc gamma-crystallins,
gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and
gammaD.";
RL Exp. Eye Res. 65:609-630(1997).
CC -!- FUNCTION: Crystallins are the dominant structural components of
the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
similar Greek key motifs.
CC -!- MISCELLANEOUS: There are six different gamma crystallins
identified in rat lens.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M19359; AAA40985.1; -.
DR EMBL; J00716; AAA40987.1; -.
DR EMBL; X00271; CAA25073.1; -.
DR PIR; A02930; CVRTG1.
DR PIR; I56381; I56381.
DR PDB; 1A5D; X-ray; A/B=1-173.
DR RGD; 2423; Cryge.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR PROSITE; PS0915; CRYSTALLIN BETAGAMMA; 4.
KW 3D-structure; Eye lens protein; Multigene family; Repeat.
FT INIT MET 0 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT STRAND 2 8
FT TURN 9 11
FT STRAND 12 18
FT STRAND 22 22
FT STRAND 26 28
FT HELIX 34 39
FT STRAND 41 47

```

FT TURN 48 50
FT STRAND 51 57
FT STRAND 60 62
FT HELIX 65 68
FT TURN 69 69
FT STRAND 77 81
FT STRAND 88 92
FT STRAND 95 97
FT STRAND 101 104
FT STRAND 108 108
FT HELIX 111 114
FT STRAND 122 127
FT STRAND 130 135
FT TURN 136 138
FT STRAND 139 145
FT STRAND 148 150
FT HELIX 153 156
FT TURN 157 157
FT STRAND 165 168
SQ SEQUENCE 173 AA; 21132 MW; 3F3200B85CB61B02 CRC64;

Query Match 70.0%; Score 773.5; DB 1; Length 173;
Best Local Similarity 77.0%; Pred. No. 5.4e-65;
Matches 134; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 7 GRIKFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYERPNOYQHGYFLRRG 66
DB 1 GKITYEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYEQPNFTGQCYFLRRG 60

QY 67 DYPDYQOMWGFNDISIRSCRLIPQHTGTFRMIRYERDDFRGQWSEITDDCPSLQDRFHLTE 126
DB 61 DYPDYQOMWGFSDSVRSCLIP-HSSSHRIRIYEREDYRGQWVEITDDCPHLQDRFHFSD 119

QY 127 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGYRYLLDWMGNAMNAKVGSLRRVMDFY 180
DB 120 PHSFHVMEGWVLYEMPYNYRGQYLLRPGYRYLLDWMGNAMNAKVGSLRRVMDFY 173
```

```

RESULT 11
CRGF_BOVIN STANDARD; PRT; 173 AA.
AC P23005;
DT 01-AUG-1991 (Rel. 19, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin F (Gamma crystallin IVA).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=89293855; PubMed=2738925;
RA White H.E., Driessen H.P.C., Slingsby C., Moss D.S., Lindley P.F.;
RT "Packing interactions in the eye-lens. Structural analysis, internal
symmetry and lattice interactions of bovine gamma IVa-crystallin.";
RL J. Mol. Biol. 207:217-235 (1989).
RN [2]
SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=98035705; PubMed=9367641; DOI=10.1006/exer.1997.0368;
RA Norledge B.V., Hay R.E., Bateman O.A., Slingsby C., Driessen H.P.C.;
RT "Towards a molecular understanding of phase separation in the lens: a
comparison of the X-ray structures of two high Tc gamma-crystallins,
gammaA and gammaB, with two low Tc gamma-crystallins, gammaB and
gammaD.";
RL Exp. Eye Res. 65:609-630 (1997).
CC -1- FUNCTION: Crystallins are the dominant structural components of
the vertebrate eye lens.
CC -1- DOMAIN: Has a two-domain beta-structure, folded into four very
similar Greek key motifs.
CC -1- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -1- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
```

```

DR PDB; 1A45; X-ray; @=1-173.
DR PDB; 1M8U; X-ray; A=1-173.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G-crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SMO0247; XTALBG; 2.
DR PROSITE; PS0915; CRYSTALLIN_BETAGAMMA; 4.
KW 3D-structure; Eye lens protein; Multigene family; Repeat.
FT INIT_MET 0 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT STRAND 2 8
FT TURN 9 11
FT STRAND 12 18
FT STRAND 22 22
FT TURN 26 28
FT STRAND 34 39
FT STRAND 41 45
FT HELIX 48 50
FT STRAND 54 57
FT STRAND 60 62
FT TURN 65 69
FT STRAND 77 81
FT STRAND 88 92
FT HELIX 95 97
FT STRAND 101 104
FT STRAND 108 108
FT TURN 112 114
FT STRAND 122 127
FT STRAND 130 133
FT STRAND 142 145
FT STRAND 148 150
FT HELIX 153 156
FT TURN 157 157
FT STRAND 165 168
SQ SEQUENCE 173 AA; 20955 MW; AC19C46CC323EC90 CRC64;

Query Match 69.9%; Score 772.5; DB 1; Length 173;
Best Local Similarity 78.7%; Pred. No. 6.7e-65;
Matches 137; Conservative 16; Mismatches 20; Indels 1; Gaps 1;

QY 7 GRIKFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYERPNOYQHGYFLRRG 66
DB 1 GKITYEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCGMWLYEQPNFTGQCYFLRRG 60

QY 67 DYPDYQOMWGFNDISIRSCRLIPQHTGTFRMIRYERDDFRGQWSEITDDCPSLQDRFHLTE 126
DB 61 DYPDYQOMWGLNDSIRSCRLIP-HTGSHRLAIYEREDYRGQWVEITDDCSSLHDFHFSE 119

QY 127 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGYRYLLDWMGNAMNAKVGSLRRVMDFY 180
DB 120 IHSFNVLEGSWVLYEMTYNYRGQYLLRPGYRYLLDWMGNAMNAKVGSLRRVMDFY 173

RESULT 12
CRGF_RAT STANDARD; PRT; 173 AA.
AC P02529;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin C (Gamma crystallin 2-1).
GN Name=Crygc;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```


RA Maglott D.R., Maltais L., Marchionni L., McKensie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reid J.J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Veraldo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Iehii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expressed by day 12 of
CC gestation. Maximum levels are found at day 30-40 followed by a
CC rapid decline.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- MISCELLANEOUS: There are six different gamma crystallins
CC identified in mouse lens.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; K02584; AAA03228.1; -;
DR EMBL; X57855; CAA40990.1; -;
DR EMBL; AK014301; BAB29256.1; -;
DR PIR; A02931; CYMSG2.
DR PIR; S26811; S26811.
DR HSSP; P02528; 1A5D.
DR MGD; MGI:89525; Cryge.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; Gcrystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALbg; 2.
DR PROSITE; PS50915; CRYSTALLIN BETAGAMMA; 4.
KW Eye lens protein; Multigene family; Repeat.
FT INIT_MET 0 0 By similarity.
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT CONFLICT 46 46 E -> Q (in Ref. 2).
FT CONFLICT 51 51 A -> T (in Ref. 1 and 2).
FT CONFLICT 90 90 K -> R (in Ref. 1).
FT CONFLICT 109 109 S -> P (in Ref. 1).
SQ SEQUENCE 173 AA; 21064 MW; 4919D9883E21F9A7 CRC64;

Query Match 69.0%; Score 762.5; DB 1; Length 173;
Best Local Similarity 75.9%; Pred. NO. 5.9e-64;
Matches 132; Conservative 22; Mismatches 19; Indels 1; Gaps 1;

QY 7 GRKKEDRGFGHYSCNDCPNLPQYFSCRNSIRVLSCWMLYERPNYQGHYFLRRG 66
DB 1 GKITFYEDRGFGHYECSTDSNLPQYFSCRNSIRVDSGCWMLYEQPNFAGCYFLRRG 60

QY 67 DYPDYQQWGMFNDIRSCLIPQHTGTFRMIRYERDDFRGQMSITDDCPSLQDRFHLTE 126

Db 61 DYPDYQQWGMFSDSVRSCLIP-HSSSHRIKIYERDYGQWVEITDDCSHLQDRFHFS 119
QY 127 VHSLLNVLEGSWVLYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMDFY 180
Db 120 FHSFHVMEGYWVLYEMPYRGQYLLRPGYRRYVDWGAMNARVGSLLRIMDFY 173

Search completed: July 25, 2005, 16:36:44
Job time : 92.6419 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:05:40 ; Search time 100.497 Seconds
(without alignments)

758.153 Million cell updates/sec

Title: US-10-030-605A-21

Perfect score: 1108

Sequence: 1 MGR1KFEDRFQGHYSCN.....PNSSVDVKLAALHHHHH 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1108	100.0	197	4	AAB46850 Bovine ga
2	1066	96.2	197	4	AAB46851 Bovine ga
3	986	89.0	198	4	AAB46848 Bovine ga
4	944	85.2	198	4	AAB46849 Bovine ga
5	769.5	69.4	173	7	ADD47108 Rat Prote
6	769.5	69.4	173	7	ADD57551 Rat Prote
7	769.5	69.4	173	7	ADD47114 Rat Prote
8	769.5	69.4	173	7	ADD47112 Rat Prote
9	769.5	69.4	173	7	ADD47118 Rat Prote
10	769.5	69.4	173	7	ADD83335 Rat Prote
11	769.5	69.4	173	7	ADD57555 Rat Prote
12	757.5	68.4	173	7	ADD47104 Rat Prote
13	727.5	65.7	174	4	ADD21006 Novel hum
14	727.5	65.7	174	4	ADD47106 Human pro
15	714.5	64.5	362	7	ADD31213 Human nov
16	710.5	64.1	173	7	ADD47110 Human pro
17	710.5	64.1	173	7	ADD83337 Human pro
18	710.5	64.1	173	7	ADD47116 Human pro
19	710.5	64.1	173	7	ADD57553 Human pro
20	705	63.6	308	4	ABG21004 Novel hum
21	680	61.4	199	4	ABG21007 Novel hum
22	604	54.5	145	4	ABG21005 Novel hum
23	335	30.2	169	8	ADQ97976 Human can
24	320	28.9	507	7	ADG31503 Human nov
25	320	28.9	616	7	ADM04410 Human pro

26	320	28.9	781	5	ABB72344	Abb72344 Murine pr
27	320	28.9	888	8	ADO44168	ADO44168 Structura
28	318.5	28.7	420	5	ABP69298	ABP69298 Human pol
29	300	27.1	205	2	AAW17523	AAW17523 Human bet
30	297	26.8	204	7	ADG63507	ADG63507 Rat prote
31	297	26.8	205	8	ADP22671	ADP22671 Golden ha
32	297	26.8	211	8	ADP22673	ADP22673 Golden ha
33	288	26.0	252	5	AAU11447	AAU11447 Human cry
34	281.5	25.4	215	2	AAW17522	AAW17522 Human bet
35	278.5	25.1	542	6	ABR41645	ABR41645 Human DIT
36	278.5	25.1	962	7	ADG29065	ADG29065 Human AIM
37	278.5	25.1	1723	6	ABG74680	ABG74680 Human CGD
38	277	25.0	250	8	ADP22669	ADP22669 Golden ha
39	276.5	25.0	1080	4	AAW00803	AAW00803 Human bon
40	276.5	25.0	1637	4	AAW00916	AAW00916 Human bon
41	276.5	25.0	2263	4	AAW79000	AAW79000 Human pro
42	257	23.2	197	4	AAW47135	AAW47135 CD1PF-16,
43	257	23.2	197	4	ABG08246	ABG08246 Novel hum
44	257	23.2	197	8	ADQ19127	ADQ19127 Human sof
45	253	22.8	97	3	AAB58947	AAB58947 Breast an

ALIGNMENTS

RESULT 1

AAAB46850
ID AAB46850 standard; protein; 197 AA.

XX AAB46850;

DT 26-APR-2001 (first entry)

DE Bovine gamma-crystalline mutant protein Mui2A-His fragment.

XX Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW Biosensor; pollution detection; pollution control; gene therapy;
KW Intracellular immunization.

XX Bos taurus.

FN DE19932688-A1.

XX 18-JAN-2001.

PF 13-JUL-1999; 99DE-01032688.

XX 13-JUL-1999; 99DE-01032688.

XX (FIED/) FIEDLER U.

XX (RUDO/) RUDOLPH R.

XX Rudolph R, Fiedler U, Boehm G, Reimann C;

XX WPI; 2001-148304/16.

XX Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.

XX Claim 11; Page 18-19; 28pp; German.

XX This invention describes a novel protein (I) with beta-'leaflet'
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene

CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;

Query Match 100.0%; Score 1108; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 5.9e-110;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRIFKEDRGFGQGHYSCNDCPNLQPYFSRCNSIRVLSCGWMLYERPNYQGHQYFLRR 60
DB 1 MGRIFKEDRGFGQGHYSCNDCPNLQPYFSRCNSIRVLSCGWMLYERPNYQGHQYFLRR 60

QY 61 GDYPDYQQWGMGFNDIRSCLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRFHLT 120
DB 61 GDYPDYQQWGMGFNDIRSCLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRFHLT 120

QY 121 EVHSLNVLEGSWVLYEMPSPYGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
DB 121 EVHSLNVLEGSWVLYEMPSPYGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180

QY 181 SSVDKLAAALEHHHHHH 197
DB 181 SSVDKLAAALEHHHHHH 197

RESULT 2
AAB46851
ID AAB46851 standard; protein; 197 AA.
XX
AC AAB46851;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline protein WT-HIS fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
DR WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The

CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene
CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;

Query Match 96.2%; Score 1066; DB 4; Length 197;
Best Local Similarity 96.4%; Pred. No. 1.9e-105;
Matches 190; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGRIFKEDRGFGQGHYSCNDCPNLQPYFSRCNSIRVLSCGWMLYERPNYQGHQYFLRR 60
DB 1 MGRIFKEDRGFGQGHYSCNDCPNLQPYFSRCNSIRVLSCGWMLYERPNYQGHQYFLRR 60

QY 61 GDYPDYQQWGMGFNDIRSCLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRFHLT 120
DB 61 GDYPDYQQWGMGFNDIRSCLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRFHLT 120

QY 121 EVHSLNVLEGSWVLYEMPSPYGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
DB 121 EVHSLNVLEGSWVLYEMPSPYGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180

QY 181 SSVDKLAAALEHHHHHH 197
DB 181 SSVDKLAAALEHHHHHH 197

RESULT 3
AAB46848
ID AAB46848 standard; protein; 198 AA.
XX
AC AAB46848;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline mutant Mui2A protein fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
DR WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 16-17; 28pp; German.

XX This invention describes a novel protein (I) with beta-'leaflet'
 CC structure having surface-exposed amino acids, present in at least two
 CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
 CC protein is altered by targeted mutagenesis so that it has new, or
 CC improved, specific binding, catalytic or fluorescent properties. The
 CC invention also describes (1) DNA (III) that encodes (I); (2) RNA (III)
 CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
 CC contain (II) or (III), or their fragments that encode a functional region
 CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
 CC and therapy, in cosmetics, bioseparation and biosensors, and for
 CC pollution detection and control, e.g. for specific targeting of gene
 CC therapy vectors and for intracellular immunization. (I) can be provided
 CC with new or improved specific antibody-like binding, catalytic or
 CC fluorescent properties, without the cost and difficulties associated with
 CC producing complete or recombinant antibodies. (I) are relatively small
 CC (20 kDa) and can be expressed with other components as multifunctional
 CC fusions. They have good stability against low pH, denaturing agents and
 CC high temperatures, conditions under which antibodies are unstable
 XX
 SQ Sequence 198 AA;

Query Match 89.0%; Score 986; DB 4; Length 198;
 Best Local Similarity 99.4%; Pred. No. 6.8e-97;
 Matches 175; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MGRIFKEDRGFGQGHYSCNSDCPNLQPYFSCNSIRVLSCGWLIERPNYQGHQYFLRR 60
 Db 6 MGRIFKEDRGFGQGHYSCNSDCPNLQPYFSCNSIRVLSCGWLIERPNYQGHQYFLRR 65
 Qy 61 GDYDPYQQWGFNDISIRSCRLIPQHTGTFRMRYERDDFRGQMSBITDDCPSLQDRFHLT 120
 Db 66 GDYDPYQQWGFNDISIRSCRLIPQHTGTFRMRYERDDFRGQMSBITDDCPSLQDRFHLT 125
 Qy 121 EVHSLNVLGSGWLVLYEMPSYRGRQYLLRPPGYRRLDGMANNKVGSLRRVMDFYA 176
 Db 126 EVHSLNVLGSGWLVLYEMPSYRGRQYLLRPPGYRRLDGMANNKVGSLRRVMDFYA 181

RESULT 4
 AAB46849
 ID AAB46849 standard; protein; 198 AA.
 AC AAB46849;
 XX
 DT 26-APR-2001 (first entry)
 XX
 XX Bovine gamma-crystalline protein fragment.
 DE
 XX Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
 KW biosensor; pollution detection; pollution control; gene therapy;
 KW intracellular immunization.
 XX
 OS Bos taurus.
 XX
 XX DE19932688-A1.
 FN
 XX 18-JAN-2001.
 PD
 XX 13-JUL-1999; 99DE-01032688.
 PF
 XX 13-JUL-1999; 99DE-01032688.
 PR
 XX (FTED/) FIEDLER U.
 PA (RUDO/) RUDOLPH R.
 XX
 XX Rudolph R, Fiedler U, Boehm G, Reimann C;
 PI WPI; 2001-148304/16.
 DR
 XX Mutated proteins having beta-leaflet structure and related nucleic acid,
 PT have new or improved properties, e.g. antibody-like specific binding or
 PT catalytic activity.

XX Claim 11; Page 17-18; 28pp; German.
 PS
 XX This invention describes a novel protein (I) with beta-'leaflet',
 CC structure having surface-exposed amino acids, present in at least two
 CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
 CC protein is altered by targeted mutagenesis so that it has new, or
 CC improved, specific binding, catalytic or fluorescent properties. The
 CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
 CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
 CC contain (II) or (III), or their fragments that encode a functional region
 CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
 CC and therapy, in cosmetics, bioseparation and biosensors, and for
 CC pollution detection and control, e.g. for specific targeting of gene
 CC therapy vectors and for intracellular immunization. (I) can be provided
 CC with new or improved specific antibody-like binding, catalytic or
 CC fluorescent properties without the cost and difficulties associated with
 CC producing complete or recombinant antibodies. (I) are relatively small
 CC (20 kDa) and can be expressed with other components as multifunctional
 CC fusions. They have good stability against low pH, denaturing agents and
 CC high temperatures, conditions under which antibodies are unstable
 XX
 SQ Sequence 198 AA;

Query Match 85.2%; Score 944; DB 4; Length 198;
 Best Local Similarity 95.5%; Pred. No. 2.1e-92;
 Matches 168; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 MGRIFKEDRGFGQGHYSCNSDCPNLQPYFSCNSIRVLSCGWLIERPNYQGHQYFLRR 60
 Db 6 MGRITFYEDRGFGQGHYSCNSDCPNLQPYFSCNSIRVDSCGWLIERPNYQGHQYFLRR 65
 Qy 61 GDYDPYQQWGFNDISIRSCRLIPQHTGTFRMRYERDDFRGQMSBITDDCPSLQDRFHLT 120
 Db 66 GDYDPYQQWGFNDISIRSCRLIPQHTGTFRMRYERDDFRGQMSBITDDCPSLQDRFHLT 125
 Qy 121 EVHSLNVLGSGWLVLYEMPSYRGRQYLLRPPGYRRLDGMANNKVGSLRRVMDFYA 176
 Db 126 EVHSLNVLGSGWLVLYEMPSYRGRQYLLRPPGYRRLDGMANNKVGSLRRVMDFYA 181

RESULT 5
 ADD47108
 ID ADD47108 standard; protein; 173 AA.
 XX
 AC ADD47108;
 XX
 DT 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein P10065, SEQ ID NO 12798.
 XX
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 OS Unidentified.
 OS
 XX WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI

DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
DE Rat Protein P10065, SEQ ID NO 12804.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
OS Unidentified.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; P10065.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 173 AA;
SQ
Query Match 69.4%; Score 769.5; DB 7; Length 173;
Best Local Similarity 77.6%; Pred. No. 8.4e-74;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;
QY 2 GR1KEDEGFGHYSCNSDCPNLQYFSCNIRVLGCGWMLYERNYQCHOYELRRG 61
DB 1 GKITVEDRFGQRCYECSSDCPNLQYFSCNIRVSDGCGWMLYERNYQGYFLRRG 60
QY 62 DYDPDQWGMFNDIRSCLIPQHTGTFMRIRYERDDFRGQMSITDDCPSLQDRPHLTE 121

Db 61 DYDPDQWGMFNDIRSCLIP-YTSSHRILYERDDYRGLVSELTDCSCHDRFRUNE 119
QY 122 VHSNLNVLGSGWLVEMPSYRGQYLLRPGYRRYLDGAMNAKVGSLRRVMDFY 175
Db 120 IYSMEVLEGSWLVEMPNYRGQYLLRPGDYRRYVDWGAMDAKVGSLRRVMDLY 173

RESULT 8
ADD47112
ID ADD47112 standard; protein; 173 AA.
XX
XX AC ADD47112;
XX
XX DT 02-DEC-2004 (revised)
XX DT 29-JAN-2004 (first entry)
XX
XX DE Rat Protein AAA40981, SEQ ID NO 12802.
XX
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX OS Rattus norvegicus.
XX OS Unidentified.
XX
XX PN WO2003016475-A2.
XX
XX PD 27-FEB-2003.
XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX
XX PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; AAA40981.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 173 AA;
SQ

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 173 AA;
SQ

	Query Match	69.4%; Score 769.5; DB 7; Length 173;
	Best Local Similarity	77.6%; Pred.No.8.4e-74;
	Matches 135; Conservative	20; Mismatches 18; Indels 1; Gaps 1
Qy	2 GRKFKEKDRGFGQHYSNCSDPCNLQPVSFSCNSIRVLSGCWMLYERNYOGHQYFLRRG	61
Dd	: : : : : : : : :	
	1 GKITFYEDRGFGQGRCYCSCSDCPNLTQVFSRCSIRVDGGCWMLYERNYOGYQYFLRRG	60
Qy	62 DYDPYQQMGGFNDIRSICRLIPQHTGTFRMRIYERDDPRGOWSEITDPCSLQRHFLUTE	121
Dd	: : : : : : : : :	
	61 DYDPYQQMGGFSDSIRSCRSP-YTSSHRIILRYERDDYRGLVSELTEDCSCHDBRFALNE	119
Qy	122 VHSINVLEGSWLVPEMPSYGRGOYLRLPGEYRYRLDWGMKANAKVGSLSRVMDFY	175
Dd	: : : : : : : : :	
	120 IYSHNVLEGSWLVPEMPNYRGROYLLRPGDYRRYHDWGMDAKVGSISLRVMIDL	173

(GEHO) GEN HOSPITAL CORP.
(FARB) BAYER AG.

WoOLF C, D'urso D, Befort K, Costigan M;
WPT; 2003-268312/26.
GENBANK; AAA40981.

New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 173 AA;

Query Match
Best Local Similarity 69.4%; Score 769.5; DB 7; Length 173;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1

Gy 2 GRKFKEKDRFGQGHYHCNSDCPNLPQFYFSCNIRVLGGCMVLYERNYQGQHYFLRRG 61
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1 GKITYEDRGFGGRCYCSDCPNLQTYSRCSNRIVDSGCCMWLVERENYQGYQLRRG 60

Gy 62 DYDDYQQMGNFNDIRSCLIPQHGTGFMRRIYERDDPRGMSEITTDPCSLQDHFHLTE 121
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 61 DYDDYQQMGFNFSDFSRSRISPSIP-YTSSHRIKLRYERDDYGLVSUTEDECCHDFRLNE 119

Gy 122 VHSNVLEGSWLVEMPYSYGROYLLRPGEYRRYLDMGANNAKVSGLSLRVMDFY 175
| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 120 IYSHMVLEGSWLVEMYNRGRQYLLRPGDYRRYHDWGAMDAKVGSLSARVDMLY 173

RESULT 12

ID ADD47104

AC ADD47104;

XZ 02-DEC-2004 (revised)

Dt 29-JAN-2004 (first entry)

Rat Protein AAA40988, SEQ ID NO 12794.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury; chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.
OS Unidentified.
XX WO2003016475-A2.
PN PD
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; AAA40988.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 173 AA;
Query Match 68.4%; Score 757.5; DB 7; Length 173;
Best Local Similarity 75.3%; Pred. No. 1.6e-72;
Matches 131; Conservative 22; Mismatches 20; Indels 1; Gaps 1;
QY 2 GRIKFKEDRGFGQHYHSCNDCPNLQPYFRCNSIRVLSCGMWLYERNYQGHQYFLRRG 61
DB 1 GRITTYEDRGFGQHRHYECSSDHSNQLQPYFRCNSIRVDSGCMWLYEQPNFTCCQYFLRRG 60
QY 62 DYDPYQQQWGMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRFHFLT 121
DB 61 DYDPYQQQWGMGFNDSIRSCRLIP-HSSSHRIYERDYGQWVEITDDCPHLQDRFHFS 119
QY 122 VHSNLVLEGSWLYEMPYSYRGQYLLRPGCEYRYLLDWGAMNAKVGSLRRVMDFY 175
DB 120 FHSFHVIEGWYLVEMPYRGQYLLRPPREYRYHDWGAMNARVGLRRIMDY 173

RESULT 13
ABG21006
ID ABG21006 standard; protein; 174 AA.
XX
AC ABG21006;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20997.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS85193.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 51365; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 174 AA;
Query Match 65.7%; Score 727.5; DB 4; Length 174;
Best Local Similarity 73.6%; Pred. No. 2.6e-69;
Matches 128; Conservative 20; Mismatches 25; Indels 1; Gaps 1;
QY 1 MGRKFKEDRGFGQHYHSCNDCPNLQPYFRCNSIRVLSCGMWLYERNYQGHQYFLRR 60
DB 1 MGKITLYEDRGFGQHRHYECSSDHPNQLPYLSCNSARVDSGCMWLYEQPNFTYGLQYFLRR 60
QY 61 GDYDPYQQQWGMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSITDDCPSLQDRFHFLT 120

Db 61 GDYADHQWGLSDSVRSCLIP-HSGSHRILYEREDYRGQMIEFTEDCSCLODRFRN 119
QY 121 EVHSLNVLEGSWVLYEMPSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMD 174
Db 120 EIHSLNVLEGSWVLYELSNYRGQYLLMPGDYRRYQDWGATNARVGSLLRVDF 173

RESULT 14
ADD47106
ID ADD47106 standard; protein; 174 AA.
XX ADD47106;
AC
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Human Protein NP_008822, SEQ ID NO 12796.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
OS Unidentified.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; NP 008822.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Example 1; Page; 1017pp; English.
XX

The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human protein (described in Table 3
of the specification) which is differentially expressed during pain.
Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 174 AA;
Query Match 65.7%; Score 727.5; DB 7; Length 174;
Best Local Similarity 73.6%; Pred. No. 2.6e-69;
Matches 128; Conservative 20; Mismatches 25; Indels 1; Gaps 1;
QY 1 MGRIFKEDRGFGQHYVYNSDCPNLQYFVSCNSIRVLSCWMLYERNYGHOYFLRR 60
Db 1 MGKITLYEDRGFGQHRHYECSSDHPNLQPYLSKCNARVDGSCWMLYEQNYSGLOFLRR 60
QY 61 GDYPDYQQWGMFNDIRSCLIPQHTGTFRRMRYERDDFRGQMSBITDDCPSLODRFHLT 120
Db 61 GDYADHQWGLSDSVRSCLIP-HSGSHRILYEREDYRGQMIEFTEDCSCLODRFRN 119
QY 121 EVHSLNVLEGSWVLYEMPSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMD 174
Db 120 EIHSLNVLEGSWVLYELSNYRGQYLLMPGDYRRYQDWGATNARVGSLLRVDF 173

RESULT 15
ADC31213
ID ADC31213 standard; protein; 362 AA.
XX
AC ADC31213;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel polypeptide sequence, SEQ ID NO:1295.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 18.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Dermanac RT;
XX
XX WPI; 2003-371981/35.
DR N-PSDB; ADC30242.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 20; SEQ ID NO 1295; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody

CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 362 AA;

Query Match 64.5%; Score 714.5; DB 7; Length 362;
Best Local Similarity 72.3%; Pred. No. 1.8e-67;
Matches 125; Conservative 21; Mismatches 26; Indels 1; Gaps 1;
QY 3 RIKFKEDRGFGHHYSCNSDCENLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGD 62
DB 191 KITFYEDRDFQGRYCNICSDCENLRYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGK 250
QY 63 YPDYQOWMGFNDISRCRLIPQHTGTFMRIRYERDDFRGOMSEITDDCPSLQDRFHLTEV 122
DB 251 YPDYQHWMLGSVSQSCRIP-HTSSHKULRYERDDYRGLMSELTDDCACVPFLRLPEI 309
QY 123 HSLNLVLEGSWVLYEMPYSYRGQYLLRPGYRRLYLDWGMNAKVGSLLRRVMDPY 175
DB 310 YSLHLVLEGCWVLYEMPENYRGQYLLRPGDYRRYHDWGGADAKVGSLLRRVTDLY 362

Search completed: July 25, 2005, 16:32:06
Job time : 100.497 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	830	74.9	187	4	US-09-949-016-9254	Sequence 9254, Ap
2	727.5	65.7	179	4	US-09-949-016-11310	Sequence 11310, A
3	532.5	48.1	192	4	US-09-949-016-11340	Sequence 11340, A
4	300	27.1	205	2	US-08-729-152-8	Sequence 8, Appl
5	296	26.7	209	4	US-09-949-016-10830	Sequence 10830, A
6	294.5	26.6	204	4	US-09-949-016-9696	Sequence 9696, Ap
7	290.5	26.2	207	4	US-09-949-016-10814	Sequence 10814, A
8	288	26.0	262	4	US-09-949-016-10815	Sequence 10815, A
9	287	25.9	235	4	US-09-949-016-8099	Sequence 8099, Ap
10	281.5	25.4	215	2	US-08-729-152-1	Sequence 1, Appl
11	100	9.0	42	1	US-08-664-449-39	Sequence 39, Appl
12	86	7.8	378	4	US-08-979-847B-121	Sequence 122, App
13	86	7.8	398	4	US-08-979-847B-121	Sequence 121, App
14	82	7.4	320	2	US-08-757-653-163	Sequence 163, App
15	82	7.4	320	2	US-08-823-516-61	Sequence 61, Appl
16	82	7.4	320	3	US-08-759-038-102	Sequence 102, App
17	82	7.4	320	3	US-08-758-314-102	Sequence 102, App
18	82	7.4	320	4	US-09-684-938-102	Sequence 102, App
19	82	7.4	320	4	US-09-308-829A-102	Sequence 102, App
20	82	7.4	320	4	US-09-940-244-61	Sequence 61, Appl
21	80.5	7.3	655	4	US-09-270-767-43477	Sequence 43477, A
22	79.5	7.2	139	4	US-09-621-976-5088	Sequence 5088, Ap
23	79.5	7.2	300	3	US-09-561-756-36	Sequence 36, Appl
24	79.5	7.2	300	3	US-09-227-721-36	Sequence 36, Appl
25	79.5	7.2	300	4	US-09-954-697-36	Sequence 36, Appl
26	79.5	7.2	315	2	US-08-484-956-91	Sequence 91, Appl
27	79.5	7.2	315	2	US-08-757-653-91	Sequence 91, Appl

```
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11310
/ LENGTH: 179
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-11310

Query Match 65.7%; Score 727.5; DB 4; Length 179;
Best Local Similarity 73.6%; Pred. No. 1.9e-70;
Matches 128; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRR 60
Db |||:|||||:|:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 65
6 MGKITLYEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRR 65
QY 61 GYPDYQQWGMFNDISIRSCRLIPQHTGTFMRIRYERDDFRGOMSEITDDCPSLQDRFHLT 120
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 124
66 GYADHQWMLGSDSVRSCLLP-HGSHRIRLYERDYRQGMIEFTEDCSCLODRFRFN 124
QY 121 EVHSLNVLEGSWLVYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMD 174
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 178
125 EIHSLNVLEGSWLVYELSNYRGQYLLMPEGDYRRYQDWGATNARVGSLLRRVIDF 178

RESULT 3
US-09-949-016-11340
/ Sequence 11340, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11340
/ LENGTH: 192
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-11340

Query Match 48.1%; Score 532.5; DB 4; Length 192;
Best Local Similarity 51.2%; Pred. No. 2.1e-49;
Matches 88; Conservative 36; Mismatches 47; Indels 1; Gaps 1;

QY 3 RIKFEDRGFGQHYHSCNDCPNLQPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRRCD 62
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 80
21 KITFYEDKFNFGRRYDCDCADFFHTYLSRCNSIKVEGTWAVYERPNFAGMYILPQGE 80
QY 63 YPDYQQWGMFNDISIRSCRLIPQHT-GTFMRIRYERDDFRGOMSEITDDCPSLQDRFHLT 121
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 140
81 YPEYQWMLGNLSCRAVHLPSGGQYKIQIFEKDFSGQMYETTEDCPSPIMEQFHRE 140
QY 122 VHSLNVLEGSWLVYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMD 173
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 173
```

```
Db 141 IHSCVKVLEGWVIFYELPNYRGQYLLDKKEYRKPIDMGAASPAVQSFRRIIVE 192

RESULT 4
US-08-729-152-8
/ Sequence 8, Application US/08729152
/ Patent No. 5871739
/ GENERAL INFORMATION:
/ APPLICANT: Inoue, Bri
/ TITLE OF INVENTION: Pharmaceutical Composition
/ NUMBER OF SEQUENCES: 47
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Wenderoth, Lind & Ponack
/ STREET: Southern Building, Suite 700, 805 Fifteenth
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb storage
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE:
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/729,152
/ FILING DATE: 11-OCT-1996
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 291993/1995
/ FILING DATE: 13-OCT-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren M. Cheek, Jr.
/ REGISTRATION NUMBER: 33,367
/ REFERENCE/DOCKET NUMBER:
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-371-8850
/ TELEFAX:
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 205 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE:
/ DESCRIPTION: protein
US-08-729-152-8

Query Match 27.1%; Score 300; DB 2; Length 205;
Best Local Similarity 37.1%; Pred. No. 2.6e-24;
Matches 63; Conservative 31; Mismatches 72; Indels 4; Gaps 3;

QY 8 EDRGFGQHYHSCNDCPNL-QPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRRGDYDDY 66
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 82
23 EQENFQGHSHELNGPCPNLKETGVEKAGSVLVQAGPVGVEQANCKGQFVFEKGEYPRW 82
QY 67 QQWMPF--NDSIRSCRLIPQHTGTFMRIRYERDDFRGOMSE-ITDDCPSLQDRFHLTVEH 123
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 142
83 DSWTSSRRTDLSLSSLRPIKVDSEHKIILYENPNFTGKKMELIDDDVPSFHAHGQEKVS 142
QY 124 SINVLGWSWLVYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLRRVMD 173
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 192
143 SVRVQSGTWVGQYGYGAGLQYLLEKGDYKDSDFGAPHPQVQSVRRIRD 192

RESULT 5
US-09-949-016-10830
/ Sequence 10830, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10830
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10830

Query Match          26.7%; Score 296; DB 4; Length 209;
Best Local Similarity 37.1%; Pred. No. 7.2e-24;
Matches 63; Conservative 31; Mismatches 72; Indels 4; Gaps 3;

QY 8 EDGFGGHYSCNSDCPNL-QPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRRGDPDY 66
Db 27 EQNFQGHSHLNGPCPNLKETGVERKAGSVLQAGPWPVGEQANCKGEQVFVEKGEYPRM 86
QY 67 QQWGMF--NDSIRSCRLIPQHTGTFMRIRYERDDFRGQMS-ITDDCPSLQDRFHLTEVH 123
Db 87 DSWTSRRDLSLRLPIKVDSEHKILLYENPNFTGKKEIIDDVPSFHAHGQEKVS 146
QY 124 SLNVLEGSWLYEMPYRGRQYLLRGEYRRYLDWGAMNAKVSLLRRVMD 173
Db 147 SVMVRSGTWGVQYPGYRGLQYLLEKGDYKDSDFGAPHPQVQSVRRIRD 196

RESULT 6
US-09-949-016-9696
; Sequence 9696, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9696
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9696

Query Match          26.6%; Score 294.5; DB 4; Length 204;
Best Local Similarity 35.5%; Pred. No. 1e-23;
Matches 65; Conservative 37; Mismatches 66; Indels 15; Gaps 7;

QY 3 RIKFKEDRFGQGHYSCNSDCPNL-QPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRRG 61
Db 21 KITIYQENFQGRMEFTSSCPNVSRSPDNVSLKVESGAWIGYEHTSFCCGQQLFRG 80
QY 62 DYPDYQWGMF--DSIRSCRLI--POHTGTFMRIRYERDDFRGQMS-ITDDCPSLQDR 115
Db 81 EYPRMDWGSNAYHTEIRLMSRPICSAHNKES-KMTIFEKENFGRQWEISDDYPSLQA 139
QY 116 R-FHLTEVHSLNVLEGSWLYEMPYRGRQYLL---RGEYRRYLDWG--AMNAKVSLL 168
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10814
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10814

Query Match          26.2%; Score 290.5; DB 4; Length 207;
Best Local Similarity 36.3%; Pred. No. 2.8e-23;
Matches 65; Conservative 37; Mismatches 60; Indels 17; Gaps 7;

QY 8 EDGFGGHYSCNSDCPNL-QPYFSRCNSIRVLSCWMLYERPNYQGHQYFLRRGDPDY 66
Db 29 DEDFGQRRHEFTASCPSVLELGFETVRSKLVSGAWVGFHAGFQGGQYILERGEYPSW 88
QY 67 QQWGMFNDISIRCL-----IPOHTGTFMRIRYERDDFRGQMS-ITDDCPSLQDR-PH 118
Db 89 DAW-GGNTAYPAERLTSFRPAACANHRDS-RLTIPEQENFLKKGELSDDYPSLQAMGWE 146
QY 119 LTEVHSLNVLEGSWLYEMPYRGRQYLL---RGEYRRYLDWG--AMNAKVSLLRV 171
Db 147 GNEVGSFHVHSGAWVCSQFPGYRGFYVLECDHHSYKHFREWGSHAPTFOVQSIRRI 205

RESULT 8
US-09-949-016-10815
; Sequence 10815, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10815
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10815
```

US-09-949-016-10815

Query Match 26.0%; Score 288; DB 4; Length 262;
Best Local Similarity 36.0%; Pred. No. 7.2e-23;
Matches 63; Conservative 31; Mismatches 77; Indels 4; Gaps 3;
QY 3 RIKFKEDRGFGQHYHSCNSDCPNLQPY-FSRNCNSIRVLSCGWMVLYERPNOYQGHQYFLRRG 61
DB 70 RLWVFELENFQARRAEFSECCSNLADRGDFRVSIIIVSAGPWVAEQSNFRGEMFILEKG 129
QY 62 DYPDQOQWGF--NDSIRSCRLIPOHTGTFRMIRYERDDFRGOMSEIT-DDCPSLQDRPH 118
DB 130 EYPRWNTWSSYSDRLMSFRPIKMDAQSHKISLPEGANFKGNTIEIQGDDAPSLWVYGF 189
QY 119 LFEVHSLNVLEGSWVLYEMPSYRGQYLLRPGYRYLLDWGAMNAKVGSLRRVMD 173
DB 190 SDRVGSVKVSGTWGVGYPGYRGYQYLLPEGDFRHNWEGAFQPMQMSLRRLRD 244

RESULT 9

US-09-949-016-8099
; Sequence 8099, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8099
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8099

Query Match 25.9%; Score 287; DB 4; Length 235;
Best Local Similarity 36.0%; Pred. No. 7.9e-23;
Matches 63; Conservative 32; Mismatches 66; Indels 14; Gaps 5;
QY 8 EDRGFGQHYHSCNSDCPNL-QPYFSRCNSIRVLSCGWMVLYERPNOYQGHQYFLRRGDPYD 66
DB 54 ELENFGKCELSAECPSTDSLEKVGSIQVESGFWLAFESRAFRGEQFVLEKGDYPRW 113
QY 67 QOQWGF--NDSIRSCRLIPOHTGTFRMIRYERDDFRGOMSEIT-DDCPSL-----QDRPH 118
DB 114 DAWNSRSDSLLSLRLPLNIDSPDKHLFENPAFSGRKWEIVDDVPSLWAHGFQDR-- 171
QY 119 LFEVHSLNVLEGSWVLYEMPSYRGQYLLRPGYRYLLDWGAMNAKVGSLRRVMD 173
DB 172 ---VASVRAINGTWGVGYEPFGYRGQYVFERGEYRHNWEGAFQPMQMSLRRLRD 223

RESULT 10

US-08-729-152-1
; Sequence 1, Application US/08729152
; Patent No. 5871739
; GENERAL INFORMATION:
; APPLICANT: Inoue, Eri
; TITLE OF INVENTION: Pharmaceutical Composition
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: Southern Building, Suite 700, 805 Fifteenth
; STREET: Street, N.W.

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,152
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 291993/1995
; FILING DATE: 13-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
US-08-729-152-1
Query Match 25.4%; Score 281.5; DB 2; Length 215;
Best Local Similarity 34.1%; Pred. No. 2.7e-22;
Matches 62; Conservative 37; Mismatches 70; Indels 13; Gaps 6;
QY 3 RIKFKEDRGFGQHYHSCNSDCPNL-QPYFSRCNSIRVLSCGWMVLYERPNOYQGHQYFLRRG 61
DB 32 KITTYDQENFQGRMEFTSSCPNVSEFSDNVRSLKVESGAWIGYEHTSFCGQQFILRRG 91
QY 62 DYPDQOQWGF--NDSIRSCR-LIPOHTGTFRMIRYERDDFRGOMSEITDDCPSLQDR 116
DB 92 EYPRWDAWGSNAVHMERLMSFRPCSNHKSXMTIFEKENFGRQWEISDDYPQLQAM 151
QY 117 -FHLTEVHSLNVLEGSWVLYEMPSYRGQYLLR-----PGEYRYLLDWG--AMNAKVGSLR 169
DB 152 GWFNNEVGSMTIQSGAWVCYHVLGYRGYQYILKCDHHEGDYKHWREWSHAQTSIQSIR 211
QY 170 RV 171
DB 212 RI 213
RESULT 11
US-08-664-449-39
; Sequence 39, Application US/08664449
; Patent No. 5766905
; GENERAL INFORMATION:
; APPLICANT: Studier, F. W.
; APPLICANT: Rosenberg, Alan H.
; TITLE OF INVENTION: Cytoplasmic Bacteriophage Display System
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET: Building 902C
; CITY: Upton
; STATE: NY
; COUNTRY: US
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Query Match 7.8%; Score 86; DB 4; Length 398;
Best Local Similarity 78.9%; Pred. No. 0.75;
Matches 15; Conservative 1; Mismatches 3; Indels

```
QY 179 NSSVDKLAALAEHHHHH 197
| : |||||
Db 380 NFKSLPKLAALAEHHHHH 398

RESULT 14
US-08-757-653-163
; Sequence 163, Application US/08757653
; Patent No. 5843669
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Lyamichev, Natasha
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,653
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-757-653-163

Query Match 7.4%; Score 82; DB 2; Length 320;
Best Local Similarity 19.2%; Pred. No. 1.5;
Matches 43; Conservative 24; Mismatches 51; Indels 106; Gaps 9;

QY 71 GFNDSIRSLRIPQHT---GTFMRIR---YERDDPRQMS-----ITDD---CP 111
| : |||||
Db 106 GSEDPFQALIKELVDLLGLARLEVPGEYADDVLASLAKAEKGEYVRILTADKDLQ 165
| : |||||
QY 112 SLQDRPHLTVHSLSNVLEGSWLVYEMPSYRGQYLLRPGYRRL----- 156
| : |||||
Db 166 LLSDRHIV--LHP-----EGYLI---TPAWLWEKYLGRPDQWADYRALTGDESDNLPGVKG 216
| : |||||
QY 157 -----DWGAMNAKVGSLRRVMDFYSDPNSSVDKL----- 186
| : |||||
Db 217 IGEKTARKLLERWGSLEALKKNDRLKPAIREKILAHMDLKLSDWLAKVRITDLPLEVD 276
| : |||||
QY 187 -----AAALEHHHHH 197
| : |||||
Db 277 AKRREDPRRLRAFLERLFFGSLHFFGLLESFKALEHHHHH 320
| : |||||

RESULT 15
US-08-823-516-61
; Sequence 61, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
```

```
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; TITLE OF INVENTION: Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-823-516-61

Query Match 7.4%; Score 82; DB 2; Length 320;
Best Local Similarity 19.2%; Pred. No. 1.5;
Matches 43; Conservative 24; Mismatches 51; Indels 106; Gaps 9;

QY 71 GFNDSIRSLRIPQHT---GTFMRIR---YERDDPRQMS-----ITDD---CP 111
| : |||||
Db 106 GSEDPFQALIKELVDLLGLARLEVPGEYADDVLASLAKAEKGEYVRILTADKDLQ 165
| : |||||
QY 112 SLQDRPHLTVHSLSNVLEGSWLVYEMPSYRGQYLLRPGYRRL----- 156
| : |||||
Db 166 LLSDRHIV--LHP-----EGYLI---TPAWLWEKYLGRPDQWADYRALTGDESDNLPGVKG 216
| : |||||
QY 157 -----DWGAMNAKVGSLRRVMDFYSDPNSSVDKL----- 186
| : |||||
Db 217 IGEKTARKLLERWGSLEALKKNDRLKPAIREKILAHMDLKLSDWLAKVRITDLPLEVD 276
| : |||||
QY 187 -----AAALEHHHHH 197
| : |||||
```

Db 277 AKRPPDRRLRAFLERLSEFGSLLEHFGLLSEPKALEHHHHH 320
|||||

Search completed: July 25, 2005, 16:39:16
Job time : 26.9561 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:36:56 ; Search time 89.5152 Seconds
(without alignments)
856.072 Million cell updates/sec

Title: US-10-030-605A-21

Perfect score: 1108
Sequence: 1 MGR1KFKEDRGFGQHYHSCN.....PNSSVVDKLAALHHHHH 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1741741 seqs, 388992284 residues

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	424.5	38.3	89	16 US-10-425-115-346968	Sequence 346968,
2	320	28.9	616	15 US-10-108-260A-3095	Sequence 3095, Ap
3	320	28.9	781	10 US-09-866-050A-668	Sequence 668, App
4	257	23.2	197	16 US-10-723-860-1946	Sequence 1946, Ap
5	253	22.8	97	9 US-09-925-298-655	Sequence 655, App
6	253	22.8	97	14 US-10-102-806-655	Sequence 655, App
7	210.5	19.0	511	15 US-10-104-047-2339	Sequence 2339, Ap
8	200	18.1	138	14 US-10-106-698-5959	Sequence 5959, Ap
9	163	14.7	83	9 US-09-864-761-4293	Sequence 4293, A
10	147	13.3	30	9 US-09-808-602-62	Sequence 62, Appl
11	123.5	11.1	826	15 US-10-395-241-18	Sequence 18, Appl

12	123	11.1	168	14	US-10-360-053-2	Sequence 2, Appli
13	123	11.1	168	14	US-10-360-053-4	Sequence 4, Appli
14	119	10.7	21	14	US-10-010-160-68	Sequence 68, Appl
15	119	10.7	23	17	US-10-495-715-70	Sequence 70, Appl
16	100.5	9.1	151	14	US-10-360-053-10	Sequence 10, Appl
17	100.5	9.1	159	14	US-10-360-053-12	Sequence 12, Appl
18	95	8.6	526	9	US-09-731-221-79	Sequence 79, Appl
19	93.5	8.4	381	14	US-10-141-531-47	Sequence 47, Appl
20	93.5	8.4	381	15	US-10-290-072-47	Sequence 47, Appl
21	93	8.4	289	17	US-10-688-745-10	Sequence 10, Appl
22	93	8.4	516	9	US-09-804-626-4	Sequence 4, Appli
23	93	8.4	519	15	US-10-359-369-34	Sequence 34, Appl
24	92	8.3	381	14	US-10-141-531-48	Sequence 48, Appl
25	92	8.3	381	14	US-10-141-531-49	Sequence 49, Appl
26	92	8.3	381	14	US-10-141-531-50	Sequence 50, Appl
27	92	8.3	381	14	US-10-141-531-51	Sequence 51, Appl
28	92	8.3	381	14	US-10-141-531-52	Sequence 52, Appl
29	92	8.3	381	14	US-10-141-531-53	Sequence 53, Appl
30	92	8.3	381	14	US-10-141-531-54	Sequence 54, Appl
31	92	8.3	381	14	US-10-141-531-56	Sequence 56, Appl
32	92	8.3	381	14	US-10-141-531-58	Sequence 58, Appl
33	92	8.3	381	15	US-10-290-072-48	Sequence 48, Appl
34	92	8.3	381	15	US-10-290-072-49	Sequence 49, Appl
35	92	8.3	381	15	US-10-290-072-50	Sequence 50, Appl
36	92	8.3	381	15	US-10-290-072-51	Sequence 51, Appl
37	92	8.3	381	15	US-10-290-072-52	Sequence 52, Appl
38	92	8.3	381	15	US-10-290-072-53	Sequence 53, Appl
39	92	8.3	381	15	US-10-290-072-54	Sequence 54, Appl
40	92	8.3	381	15	US-10-290-072-56	Sequence 56, Appl
41	92	8.3	381	15	US-10-290-072-58	Sequence 58, Appl
42	92	8.3	481	14	US-10-141-531-59	Sequence 59, Appl
43	92	8.3	481	15	US-10-290-072-59	Sequence 59, Appl
44	90.5	8.2	653	15	US-10-369-493-5789	Sequence 5789, Ap
45	89.5	8.1	351	17	US-10-899-551-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1

US-10-425-115-346968
; Sequence 346968, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 346968
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MGT4577_79599C.1.pep
US-10-425-115-346968

Query Match	38.3%	Score	424.5	DB	16	Length	89
Best Local Similarity	82.2%	Pred. No.	2.3e-36				
Matches	74	Conservative	6	Mismatches	9	Indels	1
Gaps	1						
Qy	1	MGR1KFKEDRGFGQHYHSCNDCPNLQYFRCNSIRVLSGCWMLYERPNYQGHQYFLRR	60				
Db	1	MGT4577EDRGFGQCYECSSDCPNLQYFRCNSIRVVDSCWMLYERPNYQGHQYFLRR	60				
Qy	61	GDYDYOQWCGFNDSIRSLIPQHTGTFR	90				
Db	61	GDYDYOQWLGFSYIRSLIP-HTGSHR	89				

```
RESULT 2
US-10-108-260A-3095
; Sequence 3095, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3095
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3095

Query Match      28.9%; Score 320; DB 15; Length 616;
Best Local Similarity 33.0%; Pred. No. 2.1e-24;
Matches 68; Conservative 34; Mismatches 78; Indels 26; Gaps 2;

QY      8 EDGFGQGHYYSCNSDCPNLQ-----PVFSRCNSIRVLSCGCMWLYERNYQGHQYFLRR 60
Db      129 EAPGFGQGSWEVSDIYNLQPPEDSQSPHLASVGSLSVLGGCWVGYKEGFRGHQYVLEE 188

QY      61 GDYDPYQQWGMFNDSIRSCRLIPQHTGTFRMRIYERDDPRGQMSITDDCPSLQDRFHLT 120
Db      189 GEYDWSHWGGYDELLTSLRVIRDTFGDPAVVLFEAMDFEGHGVESKALPDVELVQHGCP 248

QY      121 EVHSLNLVLEGSWVLYEMPSYRGQYLLRPGYRRYLDWGAMNAKVGSLLRRV----- 172
Db      249 STQAIHVLGSGVWVAYQEVGSGEQYVLEKGYVRNCEDWGAGNSTITLASLPVLQVGEHDIH 308

QY      173 -----DFYSDPNSSVDKLA 187
Db      309 FVSKIQLSRPRDFLGDHFSFEDDQAA 334

RESULT 3
US-09-866-050A-668
; Sequence 668, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 668
; LENGTH: 781
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-668

Query Match      28.9%; Score 320; DB 10; Length 781;
Best Local Similarity 33.3%; Pred. No. 2.9e-24;
Matches 64; Conservative 34; Mismatches 74; Indels 20; Gaps 2;

QY      8 EDGFGQGHYYSCNSDCPNLQ-----PVFSRCNSIRVLSCGCMWLYERNYQGHQYFLRR 60
Db      294 EAPGFGQGSWEVSDIYNLQPPEDSQSPQLTSGVLSRLGGCWVGYKEGFRGHQYVLEE 353

US-10-108-860-1946
; Sequence 1946, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1946
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-1946

Query Match      23.2%; Score 257; DB 16; Length 197;
Best Local Similarity 34.6%; Pred. No. 2e-18; Mismatches 66; Indels 16; Gaps 7;
Matches 62; Conservative 35; Mismatches 66; Indels 16; Gaps 7;

QY      8 EDGFGQGHYYSCNSDCPNL--QPVPSCNSIRVLSCGCMWLYERNYQGHQYFLRRGDYDP 65
Db      18 DEEDFQRRRCRLLSDCANVCERGGLPRVRSVKVGVWVAPEYDPDFQGGQFILEKGDIYR 77

QY      66 YQQWNGP-----NDSIRSCR--LIPQHTGTFRMRIYERDDPRGQMSITDDCPSLQDR-FH 118
Db      78 WSAWSGSSSHNSQLLSFRPVLCAHNDS-RVTLFEGDNFQCKFDLDVDDYPSLPSMGWA 136

QY      119 LTEVHSLNLVLEGSWVLYEMPSYRGQYLL---RPGEYRRYLDWG--AMNAKVGSLLRRV 171
Db      137 SKDVGSLKVGSGAWVAYQYGRGYQYVLEDRDRHSGEFCTYGELGTQHTGQLQSLRRV 195

RESULT 5
US-09-925-298-655
; Sequence 655, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05981
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 655
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
```


Db 713 LYSRDDNGSTWTRVND-----QEHNYS---GPTMEADPKVYGRVYLGTRGRG 757
QY 153 RRYLDW-----GAMNAKVGSG-----LRRVMDPYSDPNSSVDKLA 188
Db 758 IYVADLTNKKSNEEKSTAKANGOKGTHCYVKKEIWPWCHAVPWLISDPNSSVDKLA 817
QY 189 ALEHHHHH 197
Db 818 ALEHHHHH 826

RESULT 12
US-10-360-053-2
; Sequence 2, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Uttenthal, Lars O
; APPLICANT: Nielsen, Rasmus W
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibi
; TITLE OF INVENTION: Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-2

Query Match 11.1%; Score 123; DB 14; Length 168;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 SDPNSSVDKLAALAEHHHHH 197
Db 147 SDPNSSVDKLAALAEHHHHH 168

RESULT 13
US-10-360-053-4
; Sequence 4, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Uttenthal, Lars O
; APPLICANT: Nielsen, Rasmus W
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibi
; TITLE OF INVENTION: Fragments via Antiparallel Heterogeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-4

Query Match 11.1%; Score 123; DB 14; Length 168;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 SDPNSSVDKLAALAEHHHHH 197
Db 147 SDPNSSVDKLAALAEHHHHH 168

Query Match 11.1%; Score 123; DB 14; Length 168;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 SDPNSSVDKLAALAEHHHHH 197
Db 147 SDPNSSVDKLAALAEHHHHH 168

RESULT 14
US-10-010-160-68
; Sequence 68, Application US/10010160
; Publication No. US20030103999A1
; GENERAL INFORMATION:
; APPLICANT: Rosey, Everett L.
; APPLICANT: Strugnell, Richard A.
; APPLICANT: Good, Robert T.
; APPLICANT: King, Kendall W.
; TITLE OF INVENTION: NOVEL THERAPEUTIC COMPOSITIONS FOR
; TITLE OF INVENTION: TREATING INFECTION BY LAWSONIA SPP.
; FILE REFERENCE: DAVI110.001AUS
; CURRENT APPLICATION NUMBER: US/10/010,160
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: AU PRI381
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: US 60/249,596
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide
US-10-010-160-68

Query Match 10.7%; Score 119; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 DPNSSVDKLAALAEHHHHH 197
Db 1 DPNSSVDKLAALAEHHHHH 21

RESULT 15
US-10-495-715-70
; Sequence 70, Application US/10495715
; Publication No. US20050037355A1
; GENERAL INFORMATION:
; APPLICANT: The Secretary of State for Defence
; APPLICANT: Day, John C
; APPLICANT: Squirrel, David J
; APPLICANT: Bailey, Mark J
; APPLICANT: White, Peter J
; TITLE OF INVENTION: Signal System and Elements used therein
; FILE REFERENCE: CG/P/149/WOD
; CURRENT APPLICATION NUMBER: US/10/495,715
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: GB 0127292.1
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: GB 0205201.7
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: pET-28a-c(+) cloning/expression region
US-10-495-715-70

Query Match 10.7%; Score 119; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 DPNSSVDKLAALAEHHHHH 197
Db 1 DPNSSVDKLAALAEHHHHH 21

Query Match 10.7%; Score 119; DB 17; Length 23;
 Best Local Similarity 100.0%; Pred. No. 3.4e-05;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 DNSSSSVDKLAALAEHHHHH 197
 |||||
 Db 3 DNSSSSVDKLAALAEHHHHH 23

Search completed: July 25, 2005, 17:00:46
 Job time : 90.8485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:17:00 ; Search time 20.299 Seconds
(without alignments)
933.776 Million cell updates/sec

Title: US-10-030-605A-21
Perfect score: 1108
Sequence: 1 MGRIFKEDRGQGHYSCN.....PNSSVDKLAALAEHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	943	85.1	175	1 CYBOG	gamma-crystallin I
2	891	79.5	175	2 B24060	gamma-crystallin 1
3	884	78.0	175	1 CYMSG3	gamma-B-crystallin
4	830	74.9	174	1 CYHUG1	gamma-crystallin 1
5	783.5	70.7	174	2 D24060	gamma-crystallin 2
6	778.5	70.3	174	1 CYMSG2	gamma-crystallin 2
7	778.5	70.3	174	2 I56381	gamma-C-crystallin
8	776.5	70.1	174	2 I83432	gamma-crystallin 2
9	775.5	70.0	174	2 C24060	gamma-crystallin 1
10	774.5	69.9	174	2 A24060	gamma-crystallin 1
11	773.5	69.8	173	1 CYRTG1	gamma-crystallin 1
12	770.5	69.5	173	2 S04266	gamma-crystallin I
13	770.5	69.5	174	1 CYMSG1	gamma-crystallin 1
14	767.5	69.3	174	2 JS0596	gamma-B-crystallin
15	765.5	69.1	174	1 CYMSG4	gamma-crystallin 4
16	764.5	69.0	174	2 S26811	gamma-B-crystallin
17	762.5	68.8	174	2 A24060	gamma-crystallin 4
18	756.5	68.3	174	2 I45881	gamma-C-crystallin
19	752.5	67.9	174	2 S33526	gamma-crystallin -
20	740.5	66.8	170	1 CYRTG2	gamma-crystallin 2
21	726.5	65.6	174	1 CYHUG2	gamma-crystallin 2
22	723.5	65.3	174	2 I77413	gamma-D-crystallin
23	715.5	64.6	174	2 A26912	gamma-crystallin 5
24	713.5	64.4	157	2 B25655	gamma-crystallin I
25	686	61.9	176	2 S45015	gamma-crystallin M
26	684.5	61.8	159	2 I49618	gamma-B-crystallin
27	682.5	61.6	151	2 I49614	gamma-C-crystallin
28	656	59.2	173	2 S45016	gamma-crystallin S
29	645	58.2	175	2 JN0682	gamma3-crystallin

ALIGNMENTS

RESULT 1

CYBOG

gamma-crystallin II - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 28-Feb-1986 #sequence revision 03-Feb-1994 #text change 09-Jul-2004

C:Accession: A29655; A02928; A0263; A0890; A0815; A02928; S04265

R:Hay, R.E.; Woods, W.D.; Church, R.L.; Petras, J.M.

Biochem. Biophys. Res. Commun. 146, 332-338, 1987

A:Title: cDNA clones encoding bovine gamma-crystallins.

A:Reference number: A90133; MUID:87270760; PMID:3606621

A:Accession: A29655

A:Molecule type: mRNA

A:Residues: 1-175 <HAY>

A:Cross-references: UNIPROT:P02526; GB:M16894; NID:9162918; PIDN:AAA30476.1; PID:9162918

R:Shat, S.P.; Spector, A.

DNA 3, 287-295, 1984

A:Title: Complete nucleotide sequence of a cDNA derived from calf lens gamma-crystallin

A:Reference number: A02928; MUID:85026663; PMID:6092016

A:Accession: A02928

A:Molecule type: mRNA

A:Residues: 2-119, 'S', 121-175 <BHA>

A:Cross-references: GB:K02112; GB:X01036; NID:9162916; PIDN:AAA30475.1; PID:9162917

A:Note: Initiator Met not shown

R:Croft, L.R.

Biochem. J. 128, 961-970, 1972

A:Title: The amino acid sequence of gamma-crystallin (fraction II) from calf lens.

A:Reference number: A90263; MUID:73054483; PMID:4674126

A:Accession: A90263

A:Molecule type: protein

A:Residues: 2-17, 'Q', 19-21, 'NN', 23, 'LOP', 28-39, 'VHSL', 45-46, 'MLQ', 48-49, 'D', 51, 53-54, 56,

A:Note: portions of this sequence were assigned by composition rather than by direct se

R:Chou, S.H.; Azari, P.; Hamel, M.E.

J. Protein Chem. 7, 67-80, 1988

A:Title: Physicochemical characterization of gamma-crystallins from bovine lens-hydroxy

A:Reference number: A60890; MUID:89351571; PMID:3255364

A:Accession: A60890

A:Molecule type: protein

A:Residues: 2-26 <CHI>

R:McDermott, M.J.; Gawinowicz-Kolska, M.A.; Chiesa, R.; Spector, A.

Arch. Biochem. Biophys. 262, 609-619, 1988

A:Title: The disulfide content of calf gamma-crystallin.

A:Reference number: A60815; MUID:88208422; PMID:3364984

A:Accession: A60815

A:Molecule type: protein

A:Residues: 2-26 <MCD>

R:Blundell, T.; Lindley, P.; Miller, L.; Moss, D.; Slingsby, C.; Tickle, I.; Turnell, B.

Nature 289, 771-777, 1981

A:Title: The molecular structure and stability of the eye lens: x-ray analysis of gamma-

A:Reference number: A93247; MUID:8112311; PMID:7464942

A:Contents: annotation; X-ray crystallography, 2.6 angstroms

R:White, H.E.; Driessen, H.P.C.; Slingsby, C.; Moss, D.S.; Lindley, P.F.

J. Mol. Biol. 207, 217-235, 1989

gamma-crystallin I
gamma-2-crystallin
gamma2-crystallin
gamma-M1-1 crystal
gamma1-crystallin
gamma-M1-2 crystal
gamma4-crystallin
gamma-crystallin S
gamma-crystallin I
gamma-s-crystallin
gamma-crystallin m
gamma-crystallin m
gamma-crystallin m
gamma-crystallin M
gamma-crystallin M
gamma-crystallin M

F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;89-129/Domain: crystallin repeat <GK3>
F;130-169/Domain: crystallin repeat <GK4>

Query Match 74.9%; Score 830; DB 1; Length 175;
Best Local Similarity 81.7%; Pred. No. 4.2e-71;
Matches 143; Conservative 16; Mismatches 16; Indels 0; Gaps 0;

Qy 1 MGRIKFEDRGFGQHYYSCNSDCPNLPQYFSRCNSIRVLSCGWMLYERPNTQGQVFLRR 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 1 MGKITFYEDRAFGQRSYECTDCEPLNPQYFSRCNSIRVESGCWMIYERPNYQGQVFLRR 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 61 GDVPDYQQWMGFNDISRSCLIPHTTGTFRMRYRDRDFRGOMSEITDDCPSLQDRPHLT 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 61 GEYPDYQQWMGLSDSIRSCLLPHSGAYRMKIYDRDELRGQMSLTDCCLSVQDRPHLT 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 121 EVHSLNVLEGSWLVEPMSYRGQYLRLPGEYRRLDWGMNAKVGSLSRVMDFY 175
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 121 EIHSNLNVLGSWILYENPNYGRGYLLRPGEYRRLDWGMAPNAKVGSLSRVMDLY 175
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 5
D24060
gamma-crystallin 2-2 - rat
N;Alternate names: gamma-D-crystallin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #text_change revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: D24060; I83433; S60744
R;den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A;Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A;Reference number: A92927; MUID:87060933; PMID:3783678
A;Accession: D24060
A:Molecule type: DNA
A;Residues: 1-174 <DEN>
A;Cross-references: UNIPROT:P10067; GB:M19359; NID:g203626; PIDN:AAA40984.1; PID:g203630
R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A;Reference number: 160312; MUID:89378747; PMID:2777080
A;Accession: I83433
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-174 <RES>
A;Cross-references: GB:M19359; NID:g203626; PIDN:AAA40984.1; PID:g203630
R;Ooki, K.; Amuro, N.; Shimizu, Y.; Okazaki, T.
Biochimie 76, 398-403, 1994
A;Title: High level expression of rat gamma-D-crystallin in Escherichia coli.
A;Reference number: S60744; MUID:95151850; PMID:7849105
A;Accession: S60744
A;Status: preliminary
A:Molecule type: mRNA
A;Residues: 1-174 <OOK>
A;Cross-references: EMBL:X57169; NID:g929827; PIDN:CAA40458.1; PID:g929828
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

Query Match 70.7%; Score 783.5; DB 2; Length 174;
Best Local Similarity 78.3%; Pred. No. 1e-66;
Matches 137; Conservative 19; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MGRIKFEDRGFGQHYYSCNSDCPNLPQYFSRCNSIRVLSCGWMLYERPNTQGQVFLRR 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 1 MGKITFYEDRGFGQRHYECSTDSNLPQYFSRCNSVRVDSGCWMLYEQPNTGCVFLRR 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 61 GDVPDYQQWMGFNDISRSCLIPHTTGTFRMRYRDRDFRGOMSEITDDCPSLQDRPHLT 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 61 GDVPDYQQWMGFSDSVSRVSCRLLP-HAGSHRILRYREDYRGQVFETDDCPSLQDRPHFN 119
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 121 EVHSLNVLEGSWLVEPMSYRGQYLRLPGEYRRLDWGMNAKVGSLSRVMDFY 175
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

```
Query Match      70.3%; Score 778.5; DB 2; Length 174;
Best Local Similarity 77.1%; Pred. No. 3e-66;
Matches 135; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGFGQGHYYSCNDCPNLQPYFSCNCSIRVLSGCWMLYERPNYQGHQYFLRR 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGKITFYEDRGFGQGRHYECSTHNSNLQPYFSCNCSVRVDSGCWMLYEQNFYGCQYFLRR 60

QY 61 GDYPDYQOMMGFNDISIRSCRLIPQHTGTFRMRIYERDDFRGOMSEITDDCPSLQDRFHLT 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GDYPDYQOMMGFSDSVRSCLLIP-HSSSHRIRIYERDYRGQWVEITDDCPHLQDRFHFS 119

QY 121 EVHSLNLVLEGSWLVYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFY 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 DFHSFHVMEGYWLYEMPYNYRGQYLLRPGYRRYLDWGAMNARVGSLLRRIMDFY 174

RESULT 8
I83432
gamma-C-crystallin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C:Accession: I83432
R:Den Dunnen, J.T.; van Neck, J.W.; Creemers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A:Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A:Reference number: I60312; MUID:89378747; PMID:2777080
A:Accession: I83432
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-174 <RES>
A:Cross-references: UNIPROT:P02529; GB:M19359; NID:g203626; PIDN:AAA40983.1; PID:g203629
C:Genetics:
A:Gene: CRY-gamma-C
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication

Query Match      70.1%; Score 776.5; DB 2; Length 174;
Best Local Similarity 78.3%; Pred. No. 4.7e-66;
Matches 137; Conservative 18; Mismatches 19; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGFGQGHYYSCNDCPNLQPYFSCNCSIRVLSGCWMLYERPNYQGHQYFLRR 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGKITFYEDRGFGQGRHYECSSDCPNLQTYFSCNCSIRVDSGCWMLYERPNYQGHQYFLRR 60

QY 61 GDYPDYQOMMGFNDISIRSCRLIPQHTGTFRMRIYERDDFRGOMSEITDDCPSLQDRFHLT 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GDYPDYQOMMGFSDSVRSCLLIP-HTGSHRMRLYEKEDHKGVMWELSDCSCIQDRFHLS 119

QY 121 EVHSLNLVLEGSWLVYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFY 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 EVRSLHVLEGCWLVYEMPYNYRGQYLLRPGYRRYLDWGAVDAKAGSLRRVVLDY 174

RESULT 9
C24060
gamma-crystallin 2-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Nov-1997
C:Accession: C24060
R:den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A:Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A:Reference number: A92927; MUID:87060933; PMID:3783678
A:Accession: C24060
A:Molecule type: DNA
A:Residues: 1-174 <DEN>
A:Cross-references: GB:M19359; GB:M19354; GB:M19360; GB:M19361; GB:M19362; GB:M19363
C:Genetics:
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication
```

```
Query Match      70.0%; Score 775.5; DB 2; Length 174;
Best Local Similarity 77.7%; Pred. No. 5.8e-66;
Matches 136; Conservative 19; Mismatches 19; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGFGQGHYYSCNDCPNLQPYFSCNCSIRVLSGCWMLYERPNYQGHQYFLRR 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGKITFYEDRGFGQGRHYECSSDCPNLQTYFSCNCSVRVDSGCWMLYERPNYQGHQYFLRR 60

QY 61 GDYPDYQOMMGFNDISIRSCRLIPQHTGTFRMRIYERDDFRGOMSEITDDCPSLQDRFHLT 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GDYPDYQOMMGFSDSVRSCLLIP-HTGSHRMRLYEKEDHKGVMWELSDCSCIQDRFHLS 119

QY 121 EVHSLNLVLEGSWLVYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFY 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 EVRSLHVLEGCWLVYEMPYNYRGQYLLRPGYRRYLDWGAVDAKAGSLRRVVLDY 174

RESULT 10
A24060
gamma-crystallin 1-1 - rat
N:Alternate names: gamma-A-crystallin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A24060; I60312
R:den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A:Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A:Reference number: A92927; MUID:87060933; PMID:3783678
A:Accession: A24060
A:Molecule type: DNA
A:Residues: 1-174 <DEN>
A:Cross-references: UNIPROT:P10065; GB:M19359; NID:g203626; PIDN:AAA40981.1; PID:g203627
R:den Dunnen, J.T.; van Neck, J.W.; Creemers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A:Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A:Reference number: I60312; MUID:89378747; PMID:2777080
A:Accession: I60312
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-174 <DE2>
A:Cross-references: GB:M19359; NID:g203626; PIDN:AAA40981.1; PID:g203627
C:Genetics:
A:Gene: CRY-gamma-A
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication

Query Match      69.9%; Score 774.5; DB 2; Length 174;
Best Local Similarity 77.7%; Pred. No. 7.2e-66;
Matches 136; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGFGQGHYYSCNDCPNLQPYFSCNCSIRVLSGCWMLYERPNYQGHQYFLRR 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGKITFYEDRGFGQGRHYECSSDCPNLQTYFSCNCSIRVDSGCWMLYERPNYQGYFLRR 60

QY 61 GDYPDYQOMMGFNDISIRSCRLIPQHTGTFRMRIYERDDFRGOMSEITDDCPSLQDRFHLT 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GDYPDYQOMMGFSDSVRSCLLIP-YTSSHRIRLYERDDYRGLVSELTDSCSIHDRFRJN 119

QY 121 EVHSLNLVLEGSWLVYEMPYSYRGQYLLRPGYRRYLDWGAMNAKVGSLLRRVMDFY 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 EIYSMHVLEGSWLVYEMPYNYRGQYLLRPGYRRYLDWGAMDAKVGSLLRRVMDLY 174

RESULT 11
CYRTGI
gamma-crystallin 1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 09-Jul-2004
C:Accession: A02930
R:Moormann, R.J.M.; den Dunnen, J.T.; Bloemendal, H.; Schoenmakers, J.G.G.
Proc. Natl. Acad. Sci. U.S.A. 79, 6876-6880, 1982
```

A:Title: Extensive intragenic sequence homology in two distinct rat lens gamma-crystallin
A:Reference number: A93934; MUID:83091061; PMID:6294661
A:Accession: A02930
A:Molecule type: mRNA
A:Residues: 1-173 <MOO>
A:Cross-references: UNIPROT:P02528; GB:J00716; NID:G203634; PIDN:AAA40987.1; PID:G203635
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens
F:1-39/Domain: crystallin repeat <GK1>
F:40-82/Domain: crystallin repeat <GK2>
F:87-127/Domain: crystallin repeat <GK3>
F:128-167/Domain: crystallin repeat <GK4>

Query Match 69.8%; Score 773.5; DB 1; Length 173;
Best Local Similarity 77.0%; Pred. No. 8.9e-66;
Matches 134; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQGHYYSCNSDCPNLQPYFSCNSIRVLSCGCMWLYERPNNYQGHQYFLRRG 61
DB 1 GKITFYEDRGFGQGRHYECSTDHSLNLQPYFSCNSVRVDSGCMWLYEQNFTGCQYFLRRG 60

QY 62 DYPDYQQWGMGFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSITDDCPSLQDRPHLRT 121
DB 61 DYPDYQQWGMGFSDSVRSCLIP-HSSSHRIYEREDYRGQWVEITDDCPHLQDRPHFSD 119

QY 122 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
DB 120 FHSFHVMEGYWVLYEMPYNGRQYLLRPGEYRRYLDWGAMNAKVGSLRRIMDFY 173

RESULT 12
S04266
gamma-crystallin IV - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
A:Accession: S04266; C60890
R:White, H.E.; Driessen, H.P.C.; Slingsby, C.; Moss, D.S.; Lindley, P.F.
J. Mol. Biol. 207, 217-235, 1989
A:Title: Packing interactions in the eye-lens. Structural analysis, internal symmetry and
A:Reference number: S04265; MUID:89293855; PMID:2738925
A:Accession: S04266
A:Molecule type: protein
A:Residues: 1-173 <WHI>
R:Chiou, S.H.; Azari, P.; Himmel, M.E.
J. Protein Chem. 7, 67-80, 1988
A:Title: Physicochemical characterization of gamma-crystallins from bovine lens-hydroxy
A:Reference number: A60890; MUID:89351571; PMID:3255364
A:Accession: C60890
A:Molecule type: protein
A:Residues: 1-25 <CHI>
A:Note: 14-His, 15-Arg, 22-Cys, 23-Pro, and 24-Asp were also found and may indicate cont
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens

Query Match 69.5%; Score 770.5; DB 2; Length 173;
Best Local Similarity 77.0%; Pred. No. 1.7e-65;
Matches 134; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQGHYYSCNSDCPNLQPYFSCNSIRVLSCGCMWLYERPNNYQGHQYFLRRG 61
DB 1 GKITFYEDRGFGQGRHYECSDHSLNLQPYFSCNSIRVDSGCMWLYEQNFTGCQYFLRRG 60

QY 62 DYPDYQQWGMGFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSITDDCPSLQDRPHLRT 121
DB 61 DYPDYQQWGMGFSDSVRSCLIP-HTSSHLRIYEREDYRGQWVEITDDCSSLQDRPHFSD 119

QY 122 VHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
DB 120 IHSFHVMEGYWVLYEMPYNGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 173

RESULT 13
CYMSG1

gamma-crystallin 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
A:Accession: A02932
R:Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Treton, J.A.; Gold, R.J.M.; Tsui,
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A:Title: gamma-crystallin family of the mouse lens: structural and evolutionary relation:
A:Reference number: A94021; MUID:85088487; PMID:6096855
A:Accession: A02932
A:Molecule type: mRNA
A:Residues: 1-174 <BRE>
A:Cross-references: UNIPROT:P04342; GB:K02583; NID:gl92778; PIDN:AAA37475.1; PID:G30919;
C:Comment: there are at least seven different gamma crystallins identified in mouse len:
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens
F:2-40/Domain: crystallin repeat <GK1>
F:41-83/Domain: crystallin repeat <GK2>
F:88-128/Domain: crystallin repeat <GK3>
F:129-168/Domain: crystallin repeat <GK4>

Query Match 69.5%; Score 770.5; DB 1; Length 174;
Best Local Similarity 76.6%; Pred. No. 1.7e-65;
Matches 134; Conservative 21; Mismatches 19; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGFGQGHYYSCNSDCPNLQPYFSCNSIRVLSCGCMWLYERPNNYQGHQYFLRR 60
DB 1 MGKITFYEDRGFGQGRHYECSTDHSLNLQPYFSCNSVRVDSGCMWLYEQNFTGCQYFLRR 60

QY 61 GDYPDYQQWGMGFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSITDDCPSLQDRPHLRT 120
DB 61 GDYPDYQQWGMGFSDSVRSCLIP-HAGSHRIYEREDYRGQWVEITDDCPSLQDRPHFN 119

QY 121 EVHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
DB 120 EYISNLVLEGSWVLYDMTNYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 174

RESULT 14
JS0596
gamma-E-crystallin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Nov-1997
A:Accession: JS0596; S18835
R:Graw, J.; Coban, L.; Liebststein, A.; Werner, T.
Gene 104, 265-270, 1991
A:Title: Murine gamma E-crystallin is distinct from murine gamma 2-crystallin.
A:Reference number: JS0596; MUID:92009223; PMID:1916296
A:Accession: JS0596
A:Molecule type: DNA
A:Residues: 1-174 <GRA>
A:Cross-references: EMBL:X57855
A:Note: this sequence differs two residues from that of rat
C:Genetics:
A:Gene: gamma-E-cry
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication

Query Match 69.3%; Score 767.5; DB 2; Length 174;
Best Local Similarity 76.0%; Pred. No. 3.3e-65;
Matches 133; Conservative 22; Mismatches 19; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGFGQGHYYSCNSDCPNLQPYFSCNSIRVLSCGCMWLYERPNNYQGHQYFLRR 60
DB 1 MGKITFYEDRGFGQGRHYECSTDHSLNLQPYFSCNSVRVDSGCMWLYEQNFTGCQYFLRR 60

QY 61 GDYPDYQQWGMGFNDISIRSCRLIPQHTGTFMRRIYERDDFRGQMSITDDCPSLQDRPHLRT 120
DB 61 GDYPDYQQWGMGFSDSVRSCLIP-HSSSHRIYEREDYRGQWVEITDDCSHLQDRPHFS 119

QY 121 EVHSNLVLEGSWVLYEMPYSYRGQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
DB 120 DFHSFHVMEGYWVLYEMPYNGRQYLLRPGEYRRYLDWGAMNAKVGSLRRIMDFY 174

```
RESULT 15
CMSG4
gamma-crystallin 4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A02935; I49613
R:Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Treton, J.A.; Gold, R.J.M.; Tsui
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A:Title: gamma-crystallin family of the mouse lens: structural and evolutionary relations
A:Reference number: A94021; MUID:85088487; PMID:6096855
A:Accession: A02935
A:Molecule type: mRNA
A:Residues: 1-174 <BRE>
A:Cross-references: UNIPROT:P04345
R:Lok, S.; Tsui, L.C.; Shinohara, T.; Piatigorsky, J.; Gold, R.; Breitman, M.
Nucleic Acids Res. 12, 4517-4529, 1984
A:Title: Analysis of the mouse gamma-crystallin gene family: assignment of multiple cDNA
A:Reference number: I48353; MUID:84247318; PMID:6330674
A:Accession: I49613
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-174 <RES>
A:Cross-references: GB:K02587; NID:g192771; PIDN:AAA37473.1; PID:g387135
C:Comment: there are at least seven different gamma crystallins identified in mouse lens
C:Genetics:
A:Introns: 3/3; 84/3
C:Superfamily: beta-crystallin
C:Keywords: duplication; eye lens
F:2-40/Domain: crystallin repeat <GK1>
F:41-83/Domain: crystallin repeat <GK2>
F:88-128/Domain: crystallin repeat <GK3>
F:129-168/Domain: crystallin repeat <GK4>

Query Match          69.1%; Score 765.5; DB 1; Length 174;
Best Local Similarity 77.1%; Pred. No. 5.1e-65;
Matches 135; Conservative 19; Mismatches 20; Indels 1; Gaps 1;

QY 1 MGRIRKEDRGFGQHYHCNSDCPNLQPYFSRCNSIRVLSCGCMWLYERPNYQGHQYFLRR 60
Db 1 MGKITFYEDRGFGQCYECSSDCPNLQTYFSRCNSIRVDSGCMWLYERPNYQGYFLRR 60

QY 61 GDYPDYQQWNGFNDGIRSCRLIPQHTGTFRMIYERDDPRGQMSBITDDCPSLQDRFHLT 120
Db 61 GDYPDYQQWNGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELMDDCSCIHDRFRLH 119

QY 121 EVHSLNVLEGSWVLYEMPYGRQYLLRPGYRRLDWGAMNAKVGSLLRRVMDFY 175
Db 120 EYSHMVLEGCWVLYEMPYGRQYLLRPGYRRLDWGAMDAKVGSLLRRVMDLY 174
```

Search completed: July 25, 2005, 16:37:52
Job time : 21.299 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:15:10 ; Search time 91.1791 Seconds
(without alignments)
1106.390 Million cell updates/sec

Title: US-10-030-605A-21

Perfect score: 1108

Sequence: 1 MGR1KFEDRFGQHYSCN.....PNSSVVDKLAALBHHHHH 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	84.7	174	1	CRGB_BOVIN
2	876	79.1	174	1	CRGB_RAT
3	859	77.5	174	1	CRGB_MOUSE
4	856	77.3	175	2	Q6PHE7
5	825	74.5	174	1	CRGB_HUMAN
6	801	72.3	174	1	CRGA_BOVIN
7	778.5	70.3	173	1	CRGD_BOVIN
8	776.5	70.1	174	2	Q6PGI0
9	773.5	69.8	173	1	CRGE_RAT
10	772.5	69.7	173	1	CRGF_BOVIN
11	771.5	69.6	173	1	CRGC_RAT
12	769.5	69.4	173	1	CRGA_RAT
13	765.5	69.1	173	1	CRGD_MOUSE
14	762.5	68.8	173	1	CRGE_MOUSE
15	760.5	68.6	173	1	CRGA_MOUSE
16	757.5	68.4	173	1	CRGF_RAT
17	752.5	67.9	173	1	CRGF_MOUSE
18	751.5	67.8	173	1	CRGE_BOVIN
19	747.5	67.5	173	1	CRGD_MOUSE
20	742.5	65.2	173	1	CRGC_MOUSE
21	722.5	65.2	173	1	CRGD_HUMAN
22	721.5	65.1	173	1	CRGC_HUMAN
23	710.5	64.1	173	1	CRGA_HUMAN
24	689	62.2	174	2	Q6KRM8
25	681	61.5	175	1	CRG2_CHICO
26	671	60.6	175	2	Q6DKC9
27	654	59.0	175	2	Q7SZA6
28	651	58.8	175	1	CRBS_CHICO
29	649	58.6	175	2	Q6DJC9
30	647	58.4	175	2	Q6DER7
31	645	58.2	175	1	CRG3_XENLA

32	644	58.1	174	2	Q9PSY5	Q9psy5 xenopus lae
33	643	58.0	174	2	Q68ES5	Q68es5 xenopus lae
34	642	57.9	175	2	Q66KW2	Q66kw2 xenopus lae
35	641	57.9	175	2	Q8QFU3	Q8qfu3 cynops pyrr
36	641	57.9	189	2	Q66KR5	Q66kr5 xenopus lae
37	639	57.7	169	1	CRG2_RANTE	P02531 rana tempor
38	635	57.3	175	2	Q66L15	Q66l15 xenopus lae
39	635	57.3	177	2	Q93615	Q93615 xenopus lae
40	632	57.0	175	2	Q66KU7	Q66ku7 xenopus lae
41	631	56.9	175	1	CRG2_XENLA	Q91724 xenopus lae
42	627	56.6	177	1	CRG1_RANCA	Q91320 rana cateb
43	621	56.0	175	1	CRG1_XENLA	Q08254 xenopus lae
44	616	55.6	177	1	CRG2_RANCA	Q91321 rana cateb
45	610	55.1	172	1	CRG4_XENLA	P55941 xenopus lae

ALIGNMENTS

RESULT 1

ID	CRGB_BOVIN	STANDARD;	PRT;	174 AA.
AC	P02526;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	Gamma crystallin B (Gamma crystallin II).			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87270760; PubMed=3606621;			
RA	Hay R.E., Woods W.D., Church R.L., Petrash J.M.;			
RT	"cDNA clones encoding bovine gamma-crystallins.";			
RL	Biochem. Biophys. Res. Commun. 146:332-338(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lens;			
RX	MEDLINE=85026663; PubMed=6092016;			
RA	Bhat S.P., Spector A.;			
RT	"Complete nucleotide sequence of a cDNA derived from calf lens gamma-crystallin mRNA: presence of Alu I-like DNA sequences.";			
RL	DNA 3:287-295(1984).			
RN	[3]			
RP	SEQUENCE OF 1-25.			
RX	PubMed=3255364;			
RA	Chiou S.H., Azari P., Himmel M.E.;			
RT	"Physicochemical characterization of gamma-crystallins from bovine lens -- hydrodynamic and biochemical properties.";			
RL	J. Protein Chem. 7:67-80(1988).			
RN	[4]			
RP	SEQUENCE OF 1-25, AND DISULFIDE BOND.			
RX	PubMed=3364984;			
RA	McDermott M.J., Gawinowicz-Kolks M.A., Chiesia R., Spector A.;			
RT	"The disulfide content of calf gamma-crystallin.";			
RL	Arch. Biochem. Biophys. 262:609-619(1988).			
RN	[5]			
RP	X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS), AND REVISIONS.			
RX	MEDLINE=81123111; PubMed=7464942;			
RA	Blundell T.L., Lindley P., Miller L., Moss D., Slingby C., Tickle I.,			
RT	Turnell B., Wistow G.;			
RL	"The molecular structure and stability of the eye lens: X-ray analysis of gamma-crystallin II.";			
RN	Nature 289:771-777(1981).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RX	MEDLINE=84036195; PubMed=6631960;			
RA	Wistow G., Turnell B., Summers L., Slingby C., Moss D., Miller L.,			
RL	Lindley P., Blundell T.L.;			
RT	"X-ray analysis of the eye lens protein gamma-II crystallin at 1.9-A			


```
-!- FUNCTION: Crystallins are the dominant structural components of CC  
CC      the vertebrate eye lens.  
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very CC  
CC similar Greek key motifs.  
CC -!- MISCELLANEOUS: There are six different gamma crystallins CC  
CC identified in rat lens.  
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family. CC  
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains. CC  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration CC  
between the Swiss Institute of Bioinformatics and the EMBL outstation- CC  
the European Bioinformatics Institute. There are no restrictions on its CC  
use by non-profit institutions as long as its content is in no way CC  
modified and this statement is not removed. Usage by and for commercial CC  
entities requires a licensee agreement (see http://www.isb-sib.ch/announce/ CC  
or send an email to license@isb-sib.ch).  
-----
```

```
EMBL; M19359; AAA0982.1; -. DR  
PIR; B24060; B24060.          DR  
HSP; P02526; IAMB.           DR  
RGD; 2420; Crygb.            DR  
InterPro; IPRO01064; Crystallin. DR  
InterPro; IFP011024; G_Crystallin_SF. DR  
Prfam; PF00030; Crystall; 2.   DR  
PRINTS; PR01367; BGCYSTALLIN. DR  
SMART; SM00247; XTALbg; 2.    DR  
PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4. DR  
Eye lens protein; Multigene Family; Repeat. KW  
INIT_MET 0 0 FT  
DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.  
FT        40 82 Beta/gamma crystallin 'Greek key' 2.  
FT DOMAIN 83 87 Connecting peptide.  
FT        88 128 Beta/gamma crystallin 'Greek key' 3.  
FT DOMAIN 129 171 Beta/gamma crystallin 'Greek key' 4.  
FT CONFLICT 15 15 C-> S (in Ref. 2).  
SQ SEQUENCE 174 AA; 20957 MW; 076955AA5A94C70C CRC64;
```

```
Query Match                79.1%; Score 876; DB 1; Length 174;  
Best Local Similarity     88.5%; Pred.No. 5.2e-73;  
Matches 154; Conservative 9; Mismatches 11; Indels 0; Gaps 0;  
  
QY       2 GRKKFKEDRFQGHHYSNCNDCFNLPFYFSRCNSIRVLSCWMLYRPNYQHGYFLRRG 61  
         |.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.  
Db       1 GKTIFFEDRGFQRCECSNCLNQTYFSRCNSVRVDSGWNLIERPNYQHGYFLRRG 60  
  
QY      62 DYPDYQQMWGFNDISRSCLIPQHTGTFRMRIYERDDFRGMSEITDDCPSLQDRPHLTE 121  
         |||||:|||||.:.|..|..|..|..|..|..|..|..|..|..|..  
Db      61 DYPDYQQMWGFSDSIIRSCLIPQHSQTGRMKIYERDDFRGMSEITDDCLSLODRPHLS 120  
  
QY     122 VHSINVLEGSWLVIYMPSYKRGROYLRPEYRRLYLWGANNNAKVGLSRRWDPY 175  
         |||||:|||||.:.|..|..|..|..|..|..|..|..|..|..|..  
Db     121 IHSLNWMEGGCWLIYEMPSTYKRGYYLLRPGEYRRLYLWGANAANKVGSRFWMDPY 174
```

[illegible]


```
Db 1 GKITFYEDRAFQGRSVECTTDCNLPYFSRCNSIRVSGCWMYIERPNYQGHQYFLRRG 60
QY 62 DYPDYQWQMGFNDSIRSCRLIPQHTGTFMRRIYERDDFRGQMSBITDDCPSLQDRPHLTE 121
Db 61 EYPDYQWQMGFSDSIRSCCLIPPHSGAYRMKIYDRDELGRQMSSELTDDCLSLQDRPHLTE 120
QY 122 VHSNLVLEGSWLYENPYSRGQYLLRPGYRRLDYGAMNAKVGSLRRVMDFY 175
Db 121 IHSNLVLEGSWLYENPYSRGQYLLRPGYRRLDYGAMNAKVGSLRRVMDLY 174

RESULT 6
CREA_BOVIN STANDARD; PRT; 174 AA.
AC PG2527;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Gamma crystallin A (Gamma crystallin IVB).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=73054483; PubMed=4674126;
RA Groft L.R.;
RT "The amino acid sequence of gamma-crystallin (fraction II) from calf
RT lens.";
RL Biochem. J. 128:961-970(1972).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
DR HSP; P02526; IDS.
DR InterPro; IPR001064; Crystallin.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALBG; 2.
DR PROSITE; PS00915; CRYSTALLIN BETAGAMMA; 4.
KW Direct protein sequencing; Eye lens protein; Multigene family; Repeat.
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 87 Connecting peptide.
FT DOMAIN 88 128 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 129 171 Beta/gamma crystallin 'Greek key' 4.
SQ SEQUENCE 174 AA; 21003 MW; E39093A105E2CEEFCRC64;

Query Match 72.3%; Score 801; DB 1; Length 174;
Best Local Similarity 82.8%; Pred. No. 4.5e-66;
Matches 144; Conservative 13; Mismatches 17; Indels 0; Gaps 0;
QY 2 GRIKFEDRGFGHGYSCNDCPNLPYFSRCNSIRVLSCGWMYIERPNYQGHQYFLRRG 61
Db 1 GKITFYEDRGFGHGYSCNDCPNLPYFSRCNSIRVDSWVYQRPDIRGHQWMLQRG 60
QY 62 DYPDYQWQMGFNDSIRSCRLIPQHTGTFMRRIYERDDFRGQMSBITDDCPSLQDRPHLTE 121
Db 61 NYPQYQWQMGFSDSIRSCCLIPPHSGAYRMKIYDRDELGRQMSSELTDDCPSLQDRPHLTE 120
QY 122 VHSNLVLEGSWLYENPYSRGQYLLRPGYRRLDYGAMNAKVGSLRRVMDFY 175
Db 121 VNSRVLEGSWLYENPYSRGQYLLRPGYRRLDYGAMNAKVGSLRRVMDLY 174

RESULT 7
CREG_BOVIN STANDARD; PRT; 173 AA.
AC P08209; Q28089;
```

```
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin D (Gamma crystallin IIIB).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Lens;
RX MEDLINE=95010404; PubMed=7925695; DOI=10.1006/exer.1994.1052;
RA Hay R.E., Andley U.P., Petrash J.M.;
RT "Expression of recombinant bovine gamma B-, gamma C- and gamma D-
RT crystallins and correlation with native proteins.";
RL Exp. Eye Res. 58:573-584(1994).
RN [2]
RP SEQUENCE OF 1-156 FROM N.A.
RC TISSUE=Lens;
RX MEDLINE=87270760; PubMed=3606621;
RA Hay R.E., Woods W.D., Church R.L., Petrash J.M.;
RT "cDNA clones encoding bovine gamma-crystallins";
RL Biochem. Biophys. Res. Commun. 146:332-338(1987).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RX PubMed=15299634;
RA Chirgatzke Y.N., Driessen H.P.C., Wright G., Slingsby C., Hay R.E.,
RA Lindley P.F.;
RT "Structure of bovine eye lens gammaD (gammaIIb)-crystallin at 1.95
RT A.";
RL Acta Crystallogr. D 52:712-721(1996).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
CC EMBL; L27070; AAA98995.1; -.
CC EMBL; M16895; AAB59282.1; -.
CC PDB; 1BLP; X-ray; A/B=1-173.
CC InterPro; IPR001064; Crystallin.
CC InterPro; IPR011024; G.crystallin_SF.
CC Pfam; PF00030; Crystall; 2.
CC PRINTS; PR01367; BGCYSTALLIN.
CC SMART; SM00247; XTALBG; 2.
CC PROSITE; PS00915; CRYSTALLIN BETAGAMMA; 4.
KW 3D-structure; Direct protein sequencing; Eye lens protein;
FT INIT_MET 0 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT CONFLICT 75 75 V -> I (in Ref. 2).
FT STRAND 2 8
FT HELIX 9 11
FT STRAND 12 18
FT STRAND 22 22
FT TURN 26 28
FT STRAND 34 39
FT STRAND 42 45
FT TURN 49 50
```

```

FT STRAND 54 57
FT STRAND 60 62
FT HELIX 65 68
FT TURN 69 69
FT STRAND 77 80
FT STRAND 88 94
FT TURN 95 97
FT STRAND 98 104
FT STRAND 108 108
FT HELIX 111 114
FT TURN 115 115
FT STRAND 122 127
FT STRAND 130 135
FT TURN 136 138
FT STRAND 139 145
FT STRAND 148 150
FT HELIX 153 156
FT TURN 157 157
FT STRAND 162 162
FT STRAND 165 168
FT SEQUENCE 173 AA; 20735 MW; D9D853EB3B3F7B5F CRC64;

Query Match 70.3%; Score 778.5; DB 1; Length 173;
Best Local Similarity 78.2%; Pred. No. 5.4e-64;
Matches 136; Conservative 17; Mismatches 20; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQHYVSCNDCPNLQPYFSCNSTRVLSCGCMWLYERNYQCHQYFLRRG 61
Db 1 GKITYEDRGFGQHYECSSDHSNLQPYLGRCSVRVDSGCMWLYEQPNYLGPOYFLRRG 60

QY 62 DYPDYQOQWGFNDSTIRSCRLIPQHTGTFRMRYERDGRGOMSEITDDCPSLQDRFHLE 121
Db 61 DYPDYQOQWGLNDSVRSCRLIP-HAGSHRLRYERDYGQMIETEDCSSLQDRFHNE 119

QY 122 VHSNLVLEGSWLYEMPSPYRGQYLLRPGYRRYLDWGAMNAKVGLSRRVMDFY 175
Db 120 IHSNLVLEGSWLYELPYNRGYRQYLLRPGYRRYHDWGAMNAKVGLSRRVIDY 173

RESULT 8
CRGD RAT STANDARD; PRT; 173 AA.
AC P10067;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin D (Gamma crystallin 2-2).
GN Name=Crygd;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90236314; PubMed=1970548; DOI=10.1016/0378-1119(90)90306-C;
RA den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA Schoenmakers J.G.G.;
RT "Isolation and characterization of the rat glutamine synthetase-
RT encoding gene.";
RL Gene 87:225-232(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060933; PubMed=3783678;
RA den Dunnen J.T., Moormann R.J.M., Lubsen N.H., Schoenmakers J.G.G.;
RT "Concerted and divergent evolution within the rat gamma-crystallin
RT gene family.";
RL J. Mol. Biol. 189:37-46(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Lens;
RX MEDLINE=95151850; PubMed=7849105; DOI=10.1016/0300-9084(94)90115-5;
RA Ooki K., Amuro N., Shimizu Y., Okazaki T.;

```

```

RT "High level expression of rat gamma-D-crystallin in Escherichia
RT coli.";
RL Biochimie 76:398-403(1994).
CC -|- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -|- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -|- MISCELLANEOUS: There are six different gamma crystallins
CC identified in rat lens.
CC -|- SIMILARITY: Belongs to the beta/gamma-crystallin family;
CC -|- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19359; AAA40984.1; -.
DR EMBL; X57169; CAA40458.1; -.
DR PIR; D24060; D24060.
DR HSSP; P02528; 1A5D.
DR RGD; 2422; Crygd.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G.crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALBg; 2.
DR PROSITE; PS0915; CRYSTALLIN BETAGAMMA; 4.
KW Eye lens protein; Multigene family; Repeat.
FT INIT MET 0 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT SEQUENCE 173 AA; 20972 MW; 930B2D42D80F4117 CRC64;

Query Match 70.3%; Score 778.5; DB 1; Length 173;
Best Local Similarity 78.2%; Pred. No. 5.4e-64;
Matches 136; Conservative 19; Mismatches 18; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQHYVSCNDCPNLQPYFSCNSTRVLSCGCMWLYERNYQCHQYFLRRG 61
Db 1 GKITYEDRGFGQHYECSSDHSNLQPYLGRCSVRVDSGCMWLYEQPNYLGPOYFLRRG 60

QY 62 DYPDYQOQWGFNDSTIRSCRLIPQHTGTFRMRYERDGRGOMSEITDDCPSLQDRFHLE 121
Db 61 DYPDYQOQWGFSDSVRSCRLIP-HAGSHRLRYERDYGQWVEFTDCPSLQDRFHNE 119

QY 122 VHSNLVLEGSWLYEMPSPYRGQYLLRPGYRRYLDWGAMNAKVGLSRRVMDFY 175
Db 120 IYSLNLVLEGSWLYEMTYRGRQYLLRPGYRRYHDWGAMNARVGLSRRVMDFY 173

RESULT 9
Q6PGIO PRELIMINARY; PRT; 174 AA.
AC Q6PGIO;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Crystallin, gamma D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

```

```
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gichwood J., Schmutz J., Myers G.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smalilus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens (By similarity).
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs (By similarity).
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
DR EMBL; BC057013; AAH57013.1; -.
DR HSSP; P02528; IASD.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G.crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALB9; 2.
DR PROSITE; PS0915; CRYSTALLIN BETAGAMMA; 4.
SQ SEQUENCE 174 AA; 21118 MW; 1AFC29C0269B35A1 CRC64;

Query Match 70.1%; Score 776.5; DB 2; Length 174;
Best Local Similarity 77.1%; Pred. No. 8.3e-64;
Matches 135; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 1 MGRIFKEDRGQGHYSCNDCPNLQPYFSCNIRVLSCGWLRYERNYQCHQYFLRR 60
DB 1 MCKITFYEDRGQGHYECSTDSHNLQPYFSCNIRVLSCGWLRYEQNPFTGCQYFLRR 60

QY 61 GDYDPYQQMGFNDSIRSCRLIPQHTGTFRMIYERDDPRGQMSITDDCPSLQDRFHLT 120
DB 61 GDYDPYQQMGFSDSVRSCLLIP-HAGSHRIRLYEREYRGQWIEFTEDCPSLQDRFHN 119

QY 121 EVHSLNVLEGSVWLYEMPSYRGQVYLLRPGYRRVLDGAMNAKVGSLRRVMDFY 175
DB 120 EYSLNVLEGSVWLYDMTNYRGQVYLLRPGYRRVLDGAMNARVGSLLRRVMDFY 174

RESULT 10
CRGE_RAT
ID CRGE_RAT STANDARD; PRT; 173 AA.
AC P02528;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin E (gamma crystallin 3-1) (Gamma-2).
GN Name=Cryge;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=90236314; PubMed=1970548; DOI=10.1016/0378-1119(90)90306-C;

den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
Schoenmakers J.G.G.;
"Isolation and characterization of the rat glutamine synthetase-
encoding gene.";
Gene 87:225-232(1990).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83091061; PubMed=6294661;
RA Moormann R.J.M., den Dunnen J.T., Mulleners L., Andreoli P.,
Bloemendal H., Schoenmakers J.G.G.;
"Extensive intragenic sequence homology in two distinct rat lens
gamma-crystallin cDNAs suggests duplications of a primordial gene.";
Proc. Natl. Acad. Sci. U.S.A. 79:6876-6880(1982).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=84114867; PubMed=6319707;
RA Moormann R.J.M., den Dunnen J.T., Mulleners L., Andreoli P.,
Bloemendal H., Schoenmakers J.G.G.;
"Strict co-linearity of genetic and protein folding domains in an
intragenically duplicated rat lens gamma-crystallin gene.";
J. Mol. Biol. 171:353-368(1983).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RC TISSUE=Lens;
RX MEDLINE=98035705; PubMed=9367641; DOI=10.1006/exer.1997.0368;
RA Norledge B.V., Hay R.E., Bateman O.A., Slingsby C., Driessen H.P.C.;
"Towards a molecular understanding of phase separation in the lens: a
comparison of the X-ray structures of two high Tc gamma-crystallins,
gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and
gammaD.";
Exp. Eye Res. 65:609-630(1997).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- MISCELLANEOUS: There are six different gamma crystalline
CC identified in rat lens.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M19359; AAA40985.1; -.
CC EMBL; J00716; AAA40987.1; -.
CC EMBL; X00271; CAA25073.1; -.
CC PIR; A02930; CYRTG1.
CC PIR; I56381; I56381.
CC PDB; IASD; X-ray; A/B=1-173.
CC RGD; 2423; Cryge.
CC InterPro; IPR011024; G.crystallin_SF.
CC Pfam; PF00030; Crystall; 2.
CC PRINTS; PR01367; BGCYSTALLIN.
CC PROSITE; PS0915; CRYSTALLIN BETAGAMMA; 4.
CC 3D-structure; Eye lens protein; Multigene family; Repeat.
FW INIT MET 0 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT STRAND 2 8
FT TURN 9 11
FT STRAND 12 18
FT STRAND 22 22
FT HELIX 26 28
FT STRAND 34 39
FT STRAND 41 47
```

```

FT TURN 48 50
FT STRAND 51 52
FT STRAND 56 57
FT STRAND 60 61
FT HELIX 65 68
FT TURN 69 69
FT STRAND 77 81
FT STRAND 88 92
FT HELIX 95 97
FT STRAND 101 104
FT STRAND 108 108
FT HELIX 111 114
FT STRAND 122 127
FT STRAND 130 135
FT TURN 136 138
FT STRAND 139 145
FT STRAND 148 150
FT HELIX 153 156
FT TURN 157 157
FT STRAND 165 168
SQ SEQUENCE 173 AA; 21132 MW; 3F3200E85CB61B02 CRC64;

Query Match 69.8%; Score 773.5; DB 1; Length 173;
Best Local Similarity 77.0%; Pred. No. 1.6e-63;
Matches 134; Conservative 21; Mismatches 18; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQHYHSCNDCPNLPYFSCNSIRVLSGCWMLYERNYQGHQYFLRRG 61
Dd 1 GKITYEDRGFGQHYHSCNDCPNLPYFSCNSIRVLSGCWMLYERNYQGHQYFLRRG 60
QY 62 DYPDYQQWGMFNDISIRSCRLIPQHTGTFRMIYERDDFRGQWSEITDDCPSLQDRFHYTE 121
Dd 61 DYPDYQQWGMFSDSVRSCLIP-HSSSHRIYEREDYRGQWVEITDDCPHLQDRFHFSD 119
QY 122 VHSNLVSGSWLYEMPYSYRGQYLLRPGCEYRRYLDWGAMNAKVGSLRRVMDFY 175
Dd 120 PHSFVMEGYWLYEMPYRGQYLLRPGCEYRRYLDWGAMNAKVGSLRRVMDFY 173

```

```

RESULT 11
CRGF_BOVIN
ID CRGF_BOVIN STANDARD; PRT; 173 AA.
AC P23005;
DT 01-AUG-1991 (Rel. 19, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin F (Gamma crystallin IVA).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=89293855; PubMed=2738925;
RA White H.E., Driessen H.P.C., Slingsby C., Moss D.S., Lindley P.F.;
RT "Packing interactions in the eye-lens. Structural analysis, internal
symmetry and lattice interactions of bovine gamma Iva-crystallin.";
RL J. Mol. Biol. 207:217-235 (1989).
RN [2]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=98035705; PubMed=9367641; DOI=10.1006/exer.1997.0368;
RA Norledge B.V., Hay R.E., Bateman O.A., Slingsby C., Driessen H.P.C.;
RT "Towards a molecular understanding of phase separation in the lens: a
comparison of the X-ray structures of two high Tc gamma-crystallins,
gammaA and gammaB, with two low Tc gamma-crystallins, gammaB and
gammaD.";
RL Exp. Eye Res. 65:609-630 (1997).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma-crystallin 'Greek key' domains.

```

```

DR PDB; 1A45; X-ray; @=1-173.
DR PDB; 1M8U; X-ray; A=1-173.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G.crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SMO0247; XTALBg; 2.
DR PROSITE; PS00915; CRYSTALLIN_BETAGAMMA; 4.
KW 3D-structure; Eye lens protein; Multigene family; Repeat.
FT INIT MET 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT STRAND 2 8
FT TURN 9 11
FT STRAND 12 18
FT STRAND 22 22
FT TURN 26 28
FT STRAND 34 39
FT STRAND 41 45
FT HELIX 48 50
FT STRAND 54 57
FT STRAND 60 62
FT TURN 65 69
FT STRAND 77 81
FT STRAND 88 92
FT HELIX 95 97
FT STRAND 101 104
FT STRAND 108 108
FT TURN 112 114
FT STRAND 122 127
FT STRAND 130 133
FT STRAND 142 145
FT STRAND 148 150
FT HELIX 153 156
FT TURN 157 157
FT STRAND 165 168
SQ SEQUENCE 173 AA; 20955 MW; AC19C46CC323EC90 CRC64;

Query Match 69.7%; Score 772.5; DB 1; Length 173;
Best Local Similarity 78.7%; Pred. No. 1.9e-63;
Matches 137; Conservative 16; Mismatches 20; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQHYHSCNDCPNLPYFSCNSIRVLSGCWMLYERNYQGHQYFLRRG 61
Dd 1 GKITYEDRGFGQHYHSCNDCPNLPYFSCNSIRVLSGCWMLYERNYQGHQYFLRRG 60
QY 62 DYPDYQQWGMFNDISIRSCRLIPQHTGTFRMIYERDDFRGQWSEITDDCPSLQDRFHYTE 121
Dd 61 DYPDYQQWGMGLNDSIRSCRLIP-HTGSHRLIYEREDYRGQWVEITDCSSLHDFHFSE 119
QY 122 VHSNLVSGSWLYEMPYSYRGQYLLRPGCEYRRYLDWGAMNAKVGSLRRVMDFY 175
Dd 120 IHSFNVLEGMWLYEMTYNRGQYLLRPGDYRRYHDGATNARVGSLLRAVDYF 173

RESULT 12
CRGC_RAT
ID CRGC_RAT STANDARD; PRT; 173 AA.
AC P02529;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin C (Gamma crystallin 2-1).
GN Name=Crygc;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```



```
Best Local Similarity 77.6%; Pred. No. 3.6e-63;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQGHYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
Db 1 GKITFYEDRGFGQRCYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60
QY 62 DYPDYQOWMGFNDISIRSCRLIPQHTGTFMRMIYERDDRGOMSEITDDCPSLQDRPHLTE 121
Db 61 DYPDYQOWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFLRNE 119
QY 122 VHSLNVLGSGWLYEMPYSRGQYLLRPGCYRRYLDWGAMNAKVGLSLRRVMDFY 175
Db 120 IYSMVLEGSWLYEMPYRGQYLLRPGCYRRYLDWGAMDAKVGLSLRRVMDLY 173

RESULT 14
CRGD MOUSE
ID _CRGD MOUSE STANDARD; PRT; 173 AA.
AC P04342; O89027;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Gamma crystallin D (Gamma crystallin 1).
GN Name=Crygd;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85088487; PubMed=6096855;
RA Breitman M.L., Lok S., Wistow G., Piatigorsky J., Tretton J.A.,
RA Gold R.J.M., Tsui L.-C.;
RA "Gamma-crystallin family of the mouse lens: structural and
RT evolutionary relationships.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:7762-7766(1984).
RN [2]
RP SEQUENCE OF 3-104 FROM N.A.
RC STRAIN=102 X C3H;
RX MEDLINE=99000834; PubMed=9782080; DOI=10.1006/geno.1998.5417;
RA Klopp N., Favor J., Loester J., Lutz R.B., Neuhaeuser-Klaus A.,
RA Prescott A., Pretsch W., Quinlan R.A., Sandilands A.,
RA Vresen G.F.J.M., Graw J.;
RA "Three murine cataract mutants (Cat2) are defective in different
RT gamma-crystallin genes.";
RL Genomics 52:152-158(1998).
CC -!- FUNCTION: Crystallins are the dominant structural components of
the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
similar Greek key motifs.
CC -!- MISCELLANEOUS: There are six different gamma crystallins
identified in mouse lens.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; K02583; AAA37475.1; -.
DR EMBL; AJ224342; CAAL1908.1; -.
DR PIR; A02932; CYMSG1.
DR HSP; P08209; 1ELP.
DR MGD; MGI:88524; Crygd.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G.crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
```

```
SMART; SM00247; XTLbgl; 2.
DR PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW Eye lens protein; Multigene family; Repeat.
FT INIT_MET 0
FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
FT DOMAIN 83 86 Connecting peptide.
FT DOMAIN 87 127 Beta/gamma crystallin 'Greek key' 3.
FT DOMAIN 128 170 Beta/gamma crystallin 'Greek key' 4.
FT CONFLICT 31 H -> R (in Ref. 2).
SQ SEQUENCE 173 AA; 20877 MW; E33755FBE9636969 CRC64;

Query Match 69.1%; Score 765.5; DB 1; Length 173;
Best Local Similarity 76.4%; Pred. No. 8.5e-63;
Matches 133; Conservative 21; Mismatches 19; Indels 1; Gaps 1;

QY 2 GRIKFKEDRGFGQGHYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
Db 1 GKITFYEDRGFGQRCYECSTDSNLPYFSCNSIRVDSGCWMLYEQNFAGCQYFLRRG 60
QY 62 DYPDYQOWMGFNDISIRSCRLIPQHTGTFMRMIYERDDRGOMSEITDDCPSLQDRPHLTE 121
Db 61 DYPDYQOWMGFSDSIRSCRLIP-HAGSHRIILYEREYRGQVIEFTEDCPSLQDRPHFNE 119
QY 122 VHSLNVLGSGWLYEMPYSRGQYLLRPGCYRRYLDWGAMNAKVGLSLRRVMDFY 175
Db 120 IYSLNVLGSGWLYDMTYRGQYLLRPGCYRRYLDWGAMNAKVGLSLRRVMDFY 173

RESULT 15
CRGE MOUSE
ID _CRGE MOUSE STANDARD; PRT; 173 AA.
AC Q03740; O89028; P26999; Q9CXK5;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin E.
GN Name=Cryge;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85088487; PubMed=6096855;
RA Breitman M.L., Lok S., Wistow G., Piatigorsky J., Tretton J.A.,
RA Gold R.J.M., Tsui L.-C.;
RA "Gamma-crystallin family of the mouse lens: structural and
RT evolutionary relationships.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:7762-7766(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=(101/E1XC3H/E1)F1; TISSUE=Liver;
RX MEDLINE=92009223; PubMed=1916296; DOI=10.1016/0378-1119(91)90260-I;
RA Graw J., Coban L., Liebetstein A., Werner T.;
RA "Murine gamma E-crystallin is distinct from murine gamma 2-
RT crystallin.";
RL Gene 104:265-270(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Brusic V., Choithia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
```

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteau G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Atakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
CC -!- FUNCTION: Crystallins are the dominant structural components of
CC the vertebrate eye lens.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expressed by day 12 of
CC gestation. Maximum levels are found at day 30-40 followed by a
CC rapid decline.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC similar Greek key motifs.
CC -!- MISCELLANEOUS: There are six different gamma crystallins
CC identified in mouse lens.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02584; AAA03228.1; -;
DR EMBL; X57855; CAA40990.1; -;
DR EMBL; AK014301; BAB29256.1; -;
DR PIR; A02931; CYMSG2.
DR PIR; S26811; S26811.
DR HSP; P02528; IASD.
DR MGD; MGI:88525; Cryge.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR011024; G.crystallin_SF.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCYSTALLIN.
DR SMART; SM00247; XTALBg; 2.
DR PROSITE; PS00915; CRYSTALLIN BETAGAMMA; 4.
DR Eye lens protein; Multigene family; Repeat.
KW INIT MET 0
FT DOMAIN 1 39
FT DOMAIN 40 82
FT DOMAIN 83 86
FT DOMAIN 87 127
FT DOMAIN 128 170
FT CONFLICT 46 46
FT CONFLICT 51 51
FT CONFLICT 90 90
FT CONFLICT 109 109
FT CONFLICT 109 109
SQ SEQUENCE 173 AA; 21064 MW; 4919D9883E21F9A7 CRC64;

Query Match 68.8%; Score 762.5; DB 1; Length 173;
Best Local Similarity 75.9%; Pred. No. 1.6e-62;
Matches 132; Conservative 22; Mismatches 19; Indels 1; Gaps 1;
QY 2 GRIKFKEDRGFGCHYVCSNCPNLQPFPSRCNSIRVLSCGWMVLYERPNYQGHQYFLRRG 61
DB 1 GKITYEDRGFGCHYVCSNCPNLQPFPSRCNSIRVLSCGWMVLYEQNFAGCQYFLRRG 60
QY 62 DYPDYQWGMFNDISRCRLIPQHTGTFTFRMIYERDPRFGQWSEITDDCPSLQDRFHLE 121

Db 61 DYPDYQWGMFSDSVRSCLIP-HSSSHRIKIYEREDYRGQWVEITDDCSHLQDRFHFS 119
QY 122 VHSNLVLSGSMVLYEMPSYRGROYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175
Db 120 FHSFHVMEGYWVLYEMPYRGROYLLRPGYRRYLDWGAMNARVGSLLRRIMDPY 173

Search completed: July 25, 2005, 16:36:45
Job time : 92.1791 secs

A;Title: Packing interactions in the eye-lens. Structural analysis, internal symmetry and
 A;Reference number: S04265; MUID:89293855; PMID:2738925

A;Contents: annotation
 C;Comment: The protein has a two-domain beta-structure, folded into four very similar Gx
 C;Comment: Evidence from reference A90263 suggests an absence of disulfide bonds. Eviden
 C;Superfamily: Beta-crystallin
 F;2-40/Domain: crystallin repeat <GK1>
 F;41-83/Domain: crystallin repeat <GK2>
 F;89-129/Domain: crystallin repeat <GK3>
 F;130-169/Domain: crystallin repeat <GK4>

Query Match 88.9%; Score 990; DB 1; Length 175;
 Best Local Similarity 100.0%; Pred. No. 9.6e-85;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKITFYEDRGFGCHYECSSDCPNLPQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
 DB 1 MGKITFYEDRGFGCHYECSSDCPNLPQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
 QY 61 GDYPDYQWQMGFNDIRSCLIPQHTGTFMRIRYERDDFRGQMSSEITDDCPSLQDRFHLT 120
 DB 61 GDYPDYQWQMGFNDIRSCLIPQHTGTFMRIRYERDDFRGQMSSEITDDCPSLQDRFHLT 120
 QY 121 EVHSLNVLEGSWLYEMPSPYRGQYLLRPGYRRLDYGAMNAKVGSLLRRVMDFY 175
 DB 121 EVHSLNVLEGSWLYEMPSPYRGQYLLRPGYRRLDYGAMNAKVGSLLRRVMDFY 175

RESULT 2

B24060

gamma-crystallin 1-2 - rat
 N;Alternate names: gamma-B-crystallin
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
 C;Accession: B24060; 183431
 R;Den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
 J. Mol. Biol. 189, 37-46, 1986
 A;Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
 A;Reference number: A92927; MUID:87060933; PMID:3783678
 A;Accession: B24060

A;Molecule type: DNA
 A;Residues: 1-175 <DEN>
 A;Cross-references: UNIPROT:P10066; GB:M19359; NID:G203626; PIDN:AAA40982.1; PID:G203628
 R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
 Gene 78, 201-213, 1994

A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
 A;Reference number: 160312; MUID:89378747; PMID:2777080

A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-175 <RES>

A;Cross-references: GB:M19359; NID:G203626; PIDN:AAA40982.1; PID:G203628
 C;Genetics:
 A;Introns: 3/3; 84/3
 C;Superfamily: beta-crystallin
 C;Keywords: duplication

Query Match 83.1%; Score 925; DB 2; Length 175;
 Best Local Similarity 92.0%; Pred. No. 1.1e-78;
 Matches 161; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 1 MGKITFYEDRGFGCHYECSSDCPNLPQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
 DB 1 MGKITFYEDRGFGCHYECSSDCPNLPQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
 QY 61 GDYPDYQWQMGFNDIRSCLIPQHTGTFMRIRYERDDFRGQMSSEITDDCPSLQDRFHLT 120
 DB 61 GDYPDYQWQMGFNDIRSCLIPQHTGTFMRIRYERDDFRGQMSSEITDDCPSLQDRFHLT 120
 QY 121 EVHSLNVLEGSWLYEMPSPYRGQYLLRPGYRRLDYGAMNAKVGSLLRRVMDFY 175
 DB 121 EVHSLNVLEGSWLYEMPSPYRGQYLLRPGYRRLDYGAMNAKVGSLLRRVMDFY 175

RESULT 3

CYMSG3

gamma-B-crystallin - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 17-Mar-1987 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
 C;Accession: I48359; A02933; S33523
 R;Graw, J.; Liebstein, A.; Pietrowski, D.; Schmitt-John, T.; Werner, T.
 Gene 136, 145-156, 1993
 A;Title: Genomic sequences of murine gamma B- and gamma C-crystallin-encoding genes: pr
 A;Reference number: I48359; MUID:94123992; PMID:8293998

A;Accession: I48359
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-175 <RES>

A;Cross-references: UNIPROT:P04344; EMBL:222573; NID:G311633; PIDN:CAA80296.1; PID:G3116
 R;Bretman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Treton, J.A.; Gold, R.J.M.; Teu
 Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
 A;Title: gamma-crystallin family of the mouse lens: structural and evolutionary relation
 A;Reference number: A94021; MUID:85088487; PMID:6096855

A;Accession: A02933
 A;Molecule type: mRNA
 A;Residues: 41-150, 'M', 152-175 <BRE>

A;Comment: There are at least seven different gamma crystallins identified in mouse lens
 C;Genetics:
 A;Introns: 3/3; 84/3
 C;Superfamily: beta-crystallin
 C;Keywords: duplication; eye lens
 F;2-40/Domain: crystallin repeat <GK1>
 F;41-83/Domain: crystallin repeat <GK2>
 F;89-129/Domain: crystallin repeat <GK3>
 F;130-169/Domain: crystallin repeat <GK4>

Query Match 81.6%; Score 908; DB 1; Length 175;
 Best Local Similarity 89.7%; Pred. No. 4e-77;
 Matches 157; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGKITFYEDRGFGCHYECSSDCPNLPQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
 DB 1 MGKITFYEDRGFGCHYECSSDCPNLPQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
 QY 61 GDYPDYQWQMGFNDIRSCLIPQHTGTFMRIRYERDDFRGQMSSEITDDCPSLQDRFHLT 120
 DB 61 GDYPDYQWQMGFNDIRSCLIPQHTGTFMRIRYERDDFRGQMSSEITDDCPSLQDRFHLT 120
 QY 121 EVHSLNVLEGSWLYEMPSPYRGQYLLRPGYRRLDYGAMNAKVGSLLRRVMDFY 175
 DB 121 EVHSLNVLEGSWLYEMPSPYRGQYLLRPGYRRLDYGAMNAKVGSLLRRVMDFY 175

RESULT 4

CYHUG1

gamma-crystallin 1-2 - human

C;Species: Homo sapiens (man)
 C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
 C;Accession: A24520
 R;den Dunnen, J.T.; Moormann, R.J.M.; Cremers, F.P.M.; Schoenmakers, J.G.G.
 Gene 38, 197-204, 1985

A;Title: Two human gamma gamma-crystallin genes are linked and riddled with Alu-repeats.
 A;Reference number: A91536; MUID:86056977; PMID:4065573

A;Accession: A24520
 A;Molecule type: DNA
 A;Residues: 1-175 <DEN>
 A;Cross-references: UNIPROT:P07316; GB:M11971; NID:G181114; PIDN:AAA52113.1; PID:G181118
 C;Genetics:
 A;Gene: GDB:CRYGA; CRYGI
 A;Cross-references: GDB:119076; OMIM:123660
 A;Map position: 2q33-2q35
 A;Introns: 3/3; 84/3

C;Superfamily: beta-crystallin
 C;Keywords: duplication; eye lens
 F;2-175/Product: gamma-crystallin 1-2 #status predicted <MPT>

THIS PAGE BLANK (USPTO)

A;Title: Packing interactions in the eye-lens. Structural analysis, internal symmetry at
A;Reference number: S04265; MUID:89293855; PMID:2738925

A;Contents: annotation
C;Comment: The protein has a two-domain beta-structure, folded into four very similar G
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;89-129/Domain: crystallin repeat <GK3>
F;130-169/Domain: crystallin repeat <GK4>

Query Match 88.9%; Score 990; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 9.6e-85;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKITFYEDRGFGQHCYECSSDCPNLQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
DB 1 MGKITFYEDRGFGQHCYECSSDCPNLQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
QY 61 GDYPDYQWGMFNDISIRSCILIPQHTGTFMRRIYERDDFRGQSEITDDCPSLQDRFHLT 120
DB 61 GDYPDYQWGMFNDISIRSCILIPQHTGTFMRRIYERDDFRGQSEITDDCPSLQDRFHLT 120
QY 121 EVHSLNVLEGSWLYEMPSTGRQYLLRPGYRRLYLDWGMNAKVGSLRRVMDFY 175
DB 121 EVHSLNVLEGSWLYEMPSTGRQYLLRPGYRRLYLDWGMNAKVGSLRRVMDFY 175

RESULT 2

B24060
gamma-crystallin 1-2 - rat
N;Alternate names: gamma-B-crystallin.
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: B24060; I83431
R;Den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46; 1986

A;Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A;Reference number: A92927; MUID:87060933; PMID:3783678

A;Accession: B24060
A;Molecule type: DNA
A;Residues: 1-175 <DEN>
A;Cross-references: UNIPROT:P10066; GB:M19359; NID:g203626; PIDN:AAA40982.1; PID:g203628
R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A;Reference number: 160312; MUID:89378747; PMID:2777080

A;Accession: I83431
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA
A;Residues: 1-175 <RES>
A;Cross-references: GB:M19359; NID:g203626; PIDN:AAA40982.1; PID:g203628
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

Query Match 83.1%; Score 925; DB 2; Length 175;
Best Local Similarity 92.0%; Pred. No. 1.1e-78;
Matches 161; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 1 MGKITFYEDRGFGQHCYECSSDCPNLQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
DB 1 MGKITFYEDRGFGQHCYECSSDCPNLQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
QY 61 GDYPDYQWGMFNDISIRSCILIPQHTGTFMRRIYERDDFRGQSEITDDCPSLQDRFHLT 120
DB 61 GDYPDYQWGMFNDISIRSCILIPQHTGTFMRRIYERDDFRGQSEITDDCPSLQDRFHLT 120
QY 121 EVHSLNVLEGSWLYEMPSTGRQYLLRPGYRRLYLDWGMNAKVGSLRRVMDFY 175
DB 121 EVHSLNVLEGSWLYEMPSTGRQYLLRPGYRRLYLDWGMNAKVGSLRRVMDFY 175

RESULT 3

CYMSG3

gamma-B-crystallin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 17-Mar-1987 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48359; A02933; S33523
R;Graw, J.; Liebstein, A.; Pietrowski, D.; Schmitt-John, T.; Werner, T.
Gene 136, 145-156, 1993
A;Title: Genomic sequences of murine gamma B- and gamma C-crystallin-encoding genes: pr
A;Reference number: I48359; MUID:94123992; PMID:8293998
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-175 <RES>

A;Cross-references: UNIPROT:P04344; EMBL:Z22573; NID:g311633; PIDN:CAA80296.1; PID:g311
R;Breitman, M.L.; Lok, S.; Wistow, G.; Platiogorsky, J.; Treton, J.A.; Gold, R.J.M.; Teu
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A;Title: gamma-crystallin family of the mouse lens: structural and evolutionary relation
A;Reference number: A94021; MUID:85088487; PMID:6096855

A;Accession: A02933
A;Molecule type: mRNA
A;Residues: 41-150; M' 152-175 <BRE>
C;Comment: There are at least seven different gamma crystallins identified in mouse len
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;89-129/Domain: crystallin repeat <GK3>
F;130-169/Domain: crystallin repeat <GK4>

Query Match 81.6%; Score 908; DB 1; Length 175;
Best Local Similarity 89.7%; Pred. No. 4e-77;
Matches 157; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGKITFYEDRGFGQHCYECSSDCPNLQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
DB 1 MGKITFYEDRGFGQHCYECSSDCPNLQYFSCNSIRVDSGCWMLYERPNYQHQYFLRR 60
QY 61 GDYPDYQWGMFNDISIRSCILIPQHTGTFMRRIYERDDFRGQSEITDDCPSLQDRFHLT 120
DB 61 GDYPDYQWGMFNDISIRSCILIPQHTGTFMRRIYERDDFRGQSEITDDCPSLQDRFHLT 120
QY 121 EVHSLNVLEGSWLYEMPSTGRQYLLRPGYRRLYLDWGMNAKVGSLRRVMDFY 175
DB 121 EVHSLNVLEGSWLYEMPSTGRQYLLRPGYRRLYLDWGMNAKVGSLRRVMDFY 175

RESULT 4

CYHUG1

gamma-crystallin 1-2 - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A24520
R;Den Dunnen, J.T.; Moormann, R.J.M.; Cremers, F.P.M.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Two human gamma-crystallin genes are linked and riddled with Alu-repeats.
A;Reference number: A91536; MUID:86056977; PMID:4065573
A;Accession: A24520
A;Molecule type: DNA
A;Residues: 1-175 <DEN>

A;Cross-references: UNIPROT:P07316; GB:M1971; NID:g181114; PIDN:AAAS2113.1; PID:g18111
C;Genetics:
A;Gene: GDB:CRYG1
A;Cross-references: GDB:119076; OMIM:123660
A;Map position: 2q33-q35
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-175/Product: gamma-crystallin 1-2 #status predicted <MPT>

P;2-40/Domain: crystallin repeat <GK1>
P;41-83/Domain: crystallin repeat <GK2>
P;89-129/Domain: crystallin repeat <GK3>
P;130-169/Domain: crystallin repeat <GK4>

Query Match 77.3%; Score 860; DB 1; Length 175;
Best Local Similarity 84.0%; Pred. No. 1.2e-72;
Matches 147; Conservative 11; Indels 0; Gaps 0;
2y 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2b 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2y 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2b 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2y 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175
2b 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175

RESULT 5
gamma-C-crystallin - rat
Species: Rattus norvegicus (Norway rat)
Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
Accession: I83432
Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubben, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
Reference number: I60312; MUID:89378747; PMID:2777080
Accession: I83432
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-174 <RES>
Cross-references: UNIPROT:P02529; GB:M19359; NID:G203626; PIDN:AAA40983.1; PID:G203629
Genetics:
Gene: CRY-gamma-C
Introns: 3/3; 84/3
Superfamily: beta-crystallin
Keywords: duplication

Query Match 74.0%; Score 823.5; DB 2; Length 174;
Best Local Similarity 82.3%; Pred. No. 2.9e-69;
Matches 144; Conservative 16; Mismatches 14; Indels 1; Gaps 1;
2y 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2b 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2y 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2b 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2y 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175
2b 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175

RESULT 6
gamma-crystallin 2-1 - rat
Species: Rattus norvegicus (Norway rat)
Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Nov-1997
Accession: C24060
Den Dunnen, J.T.; Moormann, R.J.M.; Lubben, N.H.; Schoenmakers, J.G.G.
Mol. Biol. 189, 37-46, 1986
Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
Reference number: A92927; MUID:87060933; PMID:3783678
Accession: C24060
Molecule type: DNA
Residues: 1-174 <DEN>

A;Cross-references: GB:M19359; GB:M19354; GB:M19360; GB:M19361; GB:M19362; GB:M19363
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

Query Match 73.9%; Score 822.5; DB 2; Length 174;
Best Local Similarity 81.7%; Pred. No. 3.5e-69;
Matches 143; Conservative 17; Mismatches 14; Indels 1; Gaps 1;
2y 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2b 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2y 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2b 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2y 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175
2b 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175

RESULT 7
gamma-crystallin 1-1 - rat
Alternate names: gamma-A-crystallin
Species: Rattus norvegicus (Norway rat)
Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
Accession: A24060; I60312
Den Dunnen, J.T.; Moormann, R.J.M.; Lubben, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
Reference number: A92927; MUID:87060933; PMID:3783678
Accession: A24060
Molecule type: DNA
Residues: 1-174 <DEN>
Cross-references: UNIPROT:P10065; GB:M19359; NID:G203626; PIDN:AAA40981.1; PID:G203627
Genetics:
Gene: CRY-gamma-A
Introns: 3/3; 84/3
Superfamily: beta-crystallin
Keywords: duplication

Query Match 73.8%; Score 821.5; DB 2; Length 174;
Best Local Similarity 81.7%; Pred. No. 4.4e-69;
Matches 143; Conservative 18; Mismatches 13; Indels 1; Gaps 1;
2y 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2b 1 MGKITYEDRGFGQHCYECSSDCPNLPYFSCNSIRVDSGCWMLYERPNYQGHQYFLRR 60
2y 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2b 61 GDYPDYQQWGMFSDSIRSCRLIPQHTGTFRMRYERDDFRGOMSEITDDCPSLQDRPHLT 120
2y 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175
2b 121 EVHSLNVLGSGWLYEMPNYGRQYLLRPGYRRYLDWGAMNAKVGSLRRVMDPY 175

RESULT 8
gamma-crystallin 2-2 - rat
Species: Rattus norvegicus (Norway rat)
Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Nov-1997
Accession: C24060
Den Dunnen, J.T.; Moormann, R.J.M.; Lubben, N.H.; Schoenmakers, J.G.G.
Mol. Biol. 189, 37-46, 1986
Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
Reference number: A92927; MUID:87060933; PMID:3783678
Accession: C24060
Molecule type: DNA
Residues: 1-174 <DEN>

THIS PAGE BLANK (USPTO)

CC -1- FUNCTION: Crystallins are the dominant structural components of
 CC the vertebrate eye lens.
 CC -1- DOMAIN: Has a two-domain beta-structure, folded into four very
 CC similar Greek key motifs.
 CC -1- MISCELLANEOUS: There are six different gamma crystallins
 CC identified in rat lens.
 CC -1- SIMILARITY: Belongs to the beta/gamma-crystallin family.
 CC -1- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M19359; AAA40982.1; --
 CC PIR; B24060; B24060.
 CC RGD; 2420; Crygb.
 CC InterPro; IPR001064; Crystallin.
 CC InterPro; IPR011024; G crystallin_SF.
 CC Pfam; PF00030; Crystall; 2.
 CC PRINTS; PR01367; BGCYSTALLIN.
 CC SMART; SM00247; XTALB; 2.
 CC PROSITE; PS0915; CRYSTALLIN BETAGAMMA; 4.
 CC Eye lens protein; Multigene family; Repeat.
 KW INIT_MET 0
 FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
 FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
 FT DOMAIN 83 87 Connecting peptide.
 FT DOMAIN 88 128 Beta/gamma crystallin 'Greek key' 3.
 FT DOMAIN 129 171 Beta/gamma crystallin 'Greek key' 4.
 FT CONFLICT 15 15 C->S (in Ref. 2).
 SQ SEQUENCE 174 AA; 20957 MW; 076955AA5A94C70C CRC64;

Query Match 82.7%; Score 920; DB 1; Length 174;
 Best Local Similarity 92.0%; Pred. No. 6.7e-77;
 Matches 160; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 GKITFEDRGFGQCHYECSSDCNLPYFSRCNSIRVDSCWMLYERPYNQGHQYFLRRG 61
 DB 1 GKITFEDRGFGQCHYECSSDCNLPYFSRCNSIRVDSCWMLYERPYNQGHQYFLRRG 60

QY 62 DYPDYQOQWGFSDSIRSCRLIPQHTCTFRMRYIERDDFRQMSIITDDCLSLQDRPHLSE 121
 DB 61 DYPDYQOQWGFSDSIRSCRLIPQHTCTFRMRYIERDDFRQMSIITDDCLSLQDRPHLSE 120

QY 122 VHSNLVLEGSWVLYEMPSPYRGQYLLRPGYRYLDGAWNAKVGLSRVMDYF 175
 DB 121 IHSNLVNEGCWVLYEMPSPYRGQYLLRPGYRYLDGAWNAKVGLSRVMDYF 174

RESULT 3

CRGB_MOUSE STANDARD; PRT; 174 AA.
 ID CRGB_MOUSE STANDARD; PRT; 174 AA.
 AC P04344; Q61593;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Gamma crystallin B (Gamma crystallin 3).
 GN Name=Crygb;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=102 X C3H;
 RX MEDLINE=94123992; PubMed=8293998; DOI=10.1016/0378-1119(93)90458-F;
 RA Graw J., Liebsstein A., Pietrowski D., Schmitt-John T., Werner T.;
 RT "Genomic sequences of murine gamma B- and gamma C-crystallin-encoding

genes: promoter analysis and complete evolutionary pattern of mouse,
 rat and human gamma-crystallins.";
 Gene 136:145-156(1993).
 (2)
 RP SEQUENCE OF 37-174 FROM N.A.
 RX MEDLINE=85088487; PubMed=6096855;
 RA Breitman M.B., Lok S., Wistow G., Piatigorsky J., Treton J.A.,
 RA Gold R.J.M., Teui L.-C.;
 RT "Gamma-crystallin family of the mouse lens: structural and
 RT evolutionary relationships.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:7762-7766(1984).
 CC -1- FUNCTION: Crystallins are the dominant structural components of
 CC the vertebrate eye lens.
 CC -1- DOMAIN: Has a two-domain beta-structure, folded into four very
 CC similar Greek key motifs.
 CC -1- MISCELLANEOUS: There are six different gamma crystallins
 CC identified in mouse lens.
 CC -1- SIMILARITY: Belongs to the beta/gamma-crystallin family.
 CC -1- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; K02585; AAA37474.1; --
 CC EMBL; Z22573; CAA80296.1; --
 CC PIR; I48359; CYMSG3.
 CC HSP; P02526; LAMM.
 CC MGD; MGI:88522; Crygb.
 CC InterPro; IPR001064; Crystallin.
 CC InterPro; IPR011024; G crystallin_SF.
 CC Pfam; PF00030; Crystall; 2.
 CC PRINTS; PR01367; BGCYSTALLIN.
 CC SMART; SM00247; XTALB; 2.
 CC PROSITE; PS0915; CRYSTALLIN BETAGAMMA; 4.
 CC Eye lens protein; Multigene family; Repeat.
 KW INIT_MET 0
 FT DOMAIN 1 39 Beta/gamma crystallin 'Greek key' 1.
 FT DOMAIN 40 82 Beta/gamma crystallin 'Greek key' 2.
 FT DOMAIN 83 87 Connecting peptide.
 FT DOMAIN 88 128 Beta/gamma crystallin 'Greek key' 3.
 FT DOMAIN 129 171 Beta/gamma crystallin 'Greek key' 4.
 SQ SEQUENCE 174 AA; 21007 MW; 9D306049E8F7327D CRC64;

Query Match 81.1%; Score 903; DB 1; Length 174;

Best Local Similarity 89.7%; Pred. No. 2.5e-75;
Matches 156; Conservative 10; Mismatches 8; Indels 0; Gaps 0;QY 2 GKITFEDRGFGQCHYECSSDCNLPYFSRCNSIRVDSCWMLYERPYNQGHQYFLRRG 61
DB 1 GKITFEDRGFGQCHYECSSDCNLPYFSRCNSIRVDSCWMLYERPYNQGHQYFLRRG 60QY 62 DYPDYQOQWGFSDSIRSCRLIPQHTCTFRMRYIERDDFRQMSIITDDCLSLQDRPHLSE 121
DB 61 DYPDYQOQWGFSDSIRSCRLIPQHTCTFRMRYIERDDFRQMSIITDDCLSLQDRPHLSE 120QY 122 VHSNLVLEGSWVLYEMPSPYRGQYLLRPGYRYLDGAWNAKVGLSRVMDYF 175
DB 121 IHSNLVNEGCWVLYEMPSPYRGQYLLRPGYRYLDGAWNAKVGLSRVMDYF 174

RESULT 4

Q6PHP7 PRELIMINARY; PRT; 175 AA.
 ID Q6PHP7
 AC Q6PHP7; 2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Crystallin, gamma B.

THIS PAGE BLANK (USPTO)